

MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A170 025

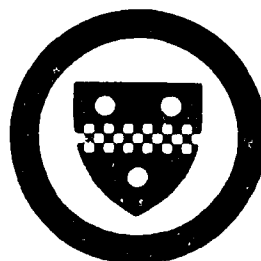
ON REDUCTION OF DIMENSIONALITY  
UNDER MULTIVARIATE REGRESSION  
AND CANONICAL CORRELATION  
MODELS

P.R. Krishnaiah  
Center for Multivariate Analysis  
University of Pittsburgh

Center for Multivariate Analysis

University of Pittsburgh

DTIC FILE COPY



DTIC  
ELECTE  
JUL 23 1986

B

DISTRIBUTION STATEMENT A  
Approved for public release  
Distribution Unlimited

| REPORT DOCUMENTATION PAGE   |                                  | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM  |
|---|----------------------------------|--|
| 1. REPORT NUMBER<br>ADSR-TR. 86-0343  | 2. GOVT ACCESSION NO.<br>A170025 | 3. RECIPIENT'S CATALOG NUMBER  |
| 4. TITLE (and Subtitle)<br>On reduction of dimensionality under multivariate regression and canonical correlation models"   |                                  | 5. TYPE OF REPORT & PERIOD COVERED<br>Technical - April 1986                       |
|   |                                  | 6. PERFORMING ORG. REPORT NUMBER<br>86-10  |
| 7. AUTHOR(s)<br>P.R. Krishnaiah   |                                  | 8. CONTRACT OR GRANT NUMBER(s)<br>F49620-85-C-0008                                 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Center for Multivariate Analysis<br>515 Thackeray Hall<br>University of Pittsburgh, Pittsburgh, PA 15260   |                                  | 10. PROGRAM ELEMENT, PROJECT, TASK<br>AREA & WORK UNIT NUMBERS<br>6.1102 F 75 2304 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Air Force Office of Scientific Research<br>Department of the Air Force<br>Bolling Air Force Base, DC 20332   |                                  | 12. REPORT DATE<br>April 1986  |
|   |                                  | 13. NUMBER OF PAGES<br>51  |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)   |                                  | 15. SECURITY CLASS. (of this report)<br>Unclassified                               |
|   |                                  | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE                                      |
| 16. DISTRIBUTION STATEMENT (of this Report)<br>Approved for public release; distribution unlimited  |                                  |  |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)  |                                  |  |
| 18. SUPPLEMENTARY NOTES   |                                  |  |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br>Contingency tables, correlated multivariate regression equations model, discriminant analysis, econometrics, likelihood ratio test, linear and structural relations, pattern recognition, random effects model, rank of canonical correlation matrix, rank of regression matrix, selection of variables, and structure of interaction.  |                                  |  |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br>In this paper, the author gives a review of the literature on various techniques for determination of the ranks of regression matrix and canonical correlation matrix. Also, methods of selection of important original variables under multivariate regression and canonical correlation models are reviewed. The methods reviewed involve not only tests of hypotheses but also model selection methods based upon information theoretic criteria. |                                  |  |

# ON REDUCTION OF DIMENSIONALITY UNDER MULTIVARIATE REGRESSION AND CANONICAL CORRELATION MODELS

P.R. Krishnaiah  
Center for Multivariate Analysis  
University of Pittsburgh

April 1986

Technical Report 86-10

Center for Multivariate Analysis  
Fifth Floor, Thackeray Hall  
University of Pittsburgh  
Pittsburgh, PA 15260

DTIC  
ELECTED  
JUL 23 1986  
B  
DESTRUCTION STATEMENT A  
Approved for public release  
Distribution Unlimited

\*Research sponsored by the Air Force Office of Scientific Research (AFSC) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)  
NOTICE OF TECHNICAL REPORTING  
This technical report is available to the public and is  
approved for release by the AFSC under contract F49620-85-C-0008.  
The report is available to the public and is approved for release by the AFSC under contract F49620-85-C-0008.  
The report is available to the public and is approved for release by the AFSC under contract F49620-85-C-0008.  
The report is available to the public and is approved for release by the AFSC under contract F49620-85-C-0008.

# ON REDUCTION OF DIMENSIONALITY UNDER MULTIVARIATE REGRESSION AND CANONICAL CORRELATION MODELS

P.R. Krishnaiah  
Center for Multivariate Analysis  
University of Pittsburgh

## ABSTRACT

In this paper, the author gives a review of the literature on various techniques for determination of the ranks of regression matrix and canonical correlation matrix. Also, methods of selection of important original variables under multivariate regression and canonical correlation models are reviewed. The methods reviewed involve not only tests of hypotheses but also model selection methods based upon information theoretic criteria.

Key words and phrases: Contingency tables, correlated multivariate regression equations model, discriminant analysis, econometrics, likelihood ratio test, linear and structural relations, pattern recognition, random effects model, rank of canonical correlation matrix, rank of regression matrix, selection of variables, and structure of interaction.

## CONTENTS

|   | Page |
|---|------|
| Abstract  | 1    |
| 1. Introduction   | 3    |
| 2. Preliminaries  | 5    |
| 3. Tests for the Rank of the Regression Matrix  | 6    |
| 4. Inference on the Rank of Regression Matrix Using Model Selection Methods                   | 13   |
| 5. Reduction of Dimensionality Under FANOVA Model   | 17   |
| 6. Rank of the Covariance Matrix of Random Effects in One-Way Components of Covariance Model  | 20   |
| 7. Selection of Original Variables Under Multivariate Regression Model                        | 23   |
| 8. Comments on Stepwise Procedures for Selection of Variables in Discriminant Analysis        | 26   |
| 9. Tests for the Rank of the Canonical Correlation Matrix                                     | 33   |
| 10. Model Selection Methods for Determination of the Rank of the Canonical Correlation Matrix | 36   |
| 11. Selection of Original Variables in Canonical Correlation Analysis                         | 38   |
| 12. Reduction of Dimensionality and the Structure of Dependence in Two-Way Contingency Table  | 41   |
| References  | 45   |



A-1

## 1. INTRODUCTION

The techniques of multivariate regression analysis and canonical correlation analysis play a very important role in the analysis of multivariate data in many disciplines. The object of this paper is to give a review of some of the work done in the literature on reduction of dimensionality in the above areas. The main emphasis of this review is on techniques for determination of the ranks of the regression matrix and canonical correlation matrix. We also review methods for selection of important original variables in the areas of multivariate regression analysis and canonical correlation analysis. This review is by no means exhaustive.

The sample regression matrix is widely used to estimate the population regression matrix under classical multivariate regression model. But, the above estimate is not the maximum likelihood estimate even when the underlying distribution is multivariate normal if the population regression matrix is not of full rank. So, it is useful to make a preliminary test to determine the rank of the regression matrix and use this information in determination of the final estimate of the regression matrix. The problem of determination of the rank of the regression matrix is also useful to determine the number of linear relations between the elements of the regression matrix. Also, the problem of determination of the number of important discriminant functions is a special case of the problem of determination of the rank of the regression matrix. The problem of determination of the rank of the canonical correlation matrix is useful in studying the relationship between two sets of variables. The number of significant canonical correlations is equivalent to the number of pairs of canonical variables which are adequate for studying the relationship between the two sets of variables. When the underlying distribution is multivariate normal, the rank of the canonical correlation matrix is equal to the rank of the regression matrix under a conditional model.

We will now mention very briefly about the importance of selection of original variables. In the area of multivariate regression analysis, it is of interest to select a small number of original variables which are adequate for prediction. Similarly, in canonical correlation analysis, it is of interest to select important original variables which are adequate to explain the relationship between two sets of variables. A brief outline of the contents of the paper is given below.

In Section 2, we give some preliminaries which are needed in the sequel. In Section 3, we first discuss the problem of determination of the number of important discriminant functions starting with the work of Fisher (1939). Then, we discuss the test procedures for the rank of the regression matrix under classical multivariate regression model. In particular, we review the work of Anderson (1951), Fujikoshi (1974), Krishnaiah, Lin and Wang (1985), Rao (1973) and Tintner (1945). In Section 4, we review the recent work of Bai, Krishnaiah and Zhao (1986a) for estimation of the rank of the regression matrix using model selection methods. These estimates are strongly consistent. In these methods, information theoretic criteria are used to select one of the various models where each model is associated with a particular rank. In Section 5, we discuss the problem of determination of the rank of the interaction matrix in two-way classification with one observation per cell. The problem of determination of the rank of the covariance matrix of the random effects in one-way multivariate random effects model is discussed in Section 6 when sample sizes of various groups are equal. The modified likelihood ratio test (LRT) procedure derived by Rao (1983) and the LRT procedure derived by Anderson (1984) and Schott and Saw (1984) for the above problem are reviewed. The model selection methods proposed by Zhao, Krishnaiah and Bai (1985a,b) recently for the above problem are also reviewed; these methods give strongly consistent estimates of the rank of the covariance matrix of the random effects for the cases when the error covariance matrix is known or unknown. A brief review of some of the methods of selection of the original variables under multivariate regression model is given in Section 7. In particular, we discuss Roy's largest root test,  $T_{\max}^2$  test, tests for additional information (Rao (1948)), and finite intersection tests (Krishnaiah (1965)). A critical review of the widely used stepwise techniques for selection of original variables in discriminant analysis is given in Section 8. We give the reasons why we should not use the above stepwise methods.

In Section 9, we review the work of Bartlett (1948), Hsu (1948a,b), Fujikoshi (1974), Lawley (1956), Krishnaiah, Lin and Wang (1985) and others on tests for the rank of the canonical correlation matrix. The recent work of Bai, Krishnaiah and Zhao (1986a) for the above problem using model selection methods is reviewed in Section 10; these methods yield strongly consistent estimates of the rank of the canonical correlation matrix. In Section 11, we review some methods of selection of original variables in canonical correlation analysis. Finally, in Section 12, we discuss the problems of reduction of

dimensionality in connection with studying the structure of dependence in two-way contingency tables. The work of Lancaster (1969), O'Neil ((1978a),(1978b),(1980)), Bhaskar Rao, Krishnaiah and Subramanyam (1985) is reviewed in the above section. The recent work of Bai, Krishnaiah and Zhao (1986b) using model selection approach is also reviewed.

## 2. PRELIMINARIES

The following notation is used throughout this paper. The transpose of a matrix is denoted by  $A'$  whereas the inverse of a square matrix is denoted by  $B^{-1}$ . The transpose of conjugate of a complex matrix  $C$  is denoted by  $C^*$ .

We now define elliptically symmetric distribution, complex multivariate normal and complex elliptically symmetric distribution. A random vector  $\underline{x} : p \times 1$  is said to have elliptically symmetric distribution if its density is of the form

$$f(\underline{x}) = |\Sigma|^{-1/2} h((\underline{x} - \underline{\mu})' \Sigma^{-1} (\underline{x} - \underline{\mu})). \quad (2.1)$$

For some details on the elliptically symmetric distribution, the reader is referred to Kelker (1970). Multivariate normal, multivariate  $t$  and multivariate Cauchy distribution are special cases of the elliptically symmetric distribution. A  $p \times 1$  random vector  $\underline{z} = \underline{x}_1 + i \underline{x}_2$  is said to be distributed as complex multivariate normal if  $\underline{x}' = (\underline{x}_1', \underline{x}_2')$  is distributed as multivariate normal with mean vector  $(\underline{\mu}_1, \underline{\mu}_2)$  and covariance matrix  $\Sigma_0$  where

$$\Sigma_0 = \begin{pmatrix} \Sigma_1 & \Sigma_2 \\ -\Sigma_2 & \Sigma_1 \end{pmatrix}$$

and  $\Sigma_1$  is of order  $p \times p$ . The complex multivariate normal distribution was considered by Wooding (1959), Goodman (1963) and others. The density function of the complex multivariate normal distribution is of the form

$$f(\underline{z}) = \pi^{-p} |\Sigma|^{-1} \exp \left[ -\frac{1}{2} (\underline{z} - \underline{\mu}_0)^* \Sigma^{-1} (\underline{z} - \underline{\mu}_0) \right] \quad (2.2)$$

where  $\Sigma = 2(\Sigma_1 - i\Sigma_2)$ ,  $\mu_0 = \mu_1 + i\mu_2$ . For a review of the literature on some multivariate distributions, the reader is referred to Krishnaiah (1976). We now define complex elliptically symmetric distribution introduced by Krishnaiah and Lin (1986). A  $p \times 1$  random vector  $\underline{z} = \underline{x}_1 + i\underline{x}_2$  is said to be distributed as complex elliptically symmetric distribution if  $\underline{x}' = (\underline{x}'_1, \underline{x}'_2)$  is distributed as elliptically symmetric distribution with density

$$f(\underline{x}) = |\Sigma_0|^{-1/2} h((\underline{x} - \underline{\mu})' \Sigma_0^{-1} (\underline{x} - \underline{\mu})) \quad (2.3)$$

where  $\underline{\mu} = (\underline{\mu}_1, \underline{\mu}_2)$ ,

$$\Sigma_0 = \begin{pmatrix} \Sigma_1 & \Sigma_2 \\ -\Sigma_2 & \Sigma_1 \end{pmatrix}$$

and  $\Sigma_1$  is of order  $p \times p$ . The density of  $\underline{z}$  is of the form

$$g(\underline{z}) = |\Sigma|^{-1} h_0((\underline{z} - \underline{\mu}_0)' \Sigma^{-1} (\underline{z} - \underline{\mu}_0)) \quad (2.4)$$

where  $\Sigma = 2(\Sigma_1 - i\Sigma_2)$  and  $\underline{\mu}_0 = \underline{\mu}_1 + i\underline{\mu}_2$ . Complex multivariate normal and complex multivariate  $t$  distributions are special cases of complex elliptically symmetric distribution.

### 3. TESTS FOR THE RANK OF THE REGRESSION MATRIX

In this section we first discuss procedures for testing the hypothesis on the number of significant discriminant functions since this is a special case of the problem of testing for the rank of regression matrix.

Let  $\underline{x}_1, \dots, \underline{x}_k$  be distributed independently as multivariate normal with mean vectors  $\underline{\mu}_1, \dots, \underline{\mu}_k$  and a common covariance matrix  $\Sigma$ . Also, let  $\underline{x}_{ij}$  ( $j=1, 2, \dots, n_i$ ) denote  $j$ -th independent observation on  $\underline{x}_i$ . Then, the between group sums of squares and cross products (SP) matrix is given by

$$S_b = \sum_{i=1}^k n_i (\bar{x}_{i\cdot} - \bar{x})(\bar{x}_{i\cdot} - \bar{x})' \quad (3.1)$$

whereas the within group SP matrix is given by

$$S_w = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})(x_{ij} - \bar{x})' \quad (3.2)$$

where

$$n\bar{x}_{i\cdot} = \sum_{j=1}^{n_i} x_{ij}, \quad n\bar{x}_{\cdot} = \sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}$$

and  $n = n_1 + \dots + n_k$ . Now, let  $\lambda_1 \geq \dots \geq \lambda_p$  denote the eigenvalues of  $S_b S_w^{-1}$ . Also, let

$$\Omega = \sum_{i=1}^k n_i (\bar{\mu}_i - \bar{\mu})(\bar{\mu}_i - \bar{\mu})' \quad (3.3)$$

where  $n\bar{\mu} = (n_1\bar{\mu}_1 + \dots + n_k\bar{\mu}_k)$ . The rank of  $\Omega$  is equivalent to the number of significant discriminant functions. Fisher (1939) proposed to use  $T_1 = (\lambda_{r+1} + \dots + \lambda_s)$  as a test statistic for testing the hypothesis that the rank of  $\Omega$  is equal to  $r$  where  $s = \min(p, k-1)$ . In general, we can use suitable functions  $\psi(\lambda_{r+1}, \dots, \lambda_s)$  of  $\lambda_{r+1}, \dots, \lambda_s$  to test for the rank of  $\Omega \Sigma^{-1}$ . For example,  $\psi(\lambda_{r+1}, \dots, \lambda_s)$  may be  $\lambda_{r+1}$ . But the distributions of these statistics involve  $\lambda_1, \dots, \lambda_r$  as nuisance parameters where  $\lambda_1 \geq \dots \geq \lambda_p$  are the eigenvalues of  $\Omega \Sigma^{-1}/n$ .

We now discuss the asymptotic joint distribution of the eigenvalues of  $S_b S_w^{-1}$  derived by Bai, Krishnaiah and Liang (1984) since it is useful in implementation of some test procedures for determination of the rank of  $\Omega$  under certain conditions. For each  $i$ , let  $x_{i1}, \dots, x_{in_i}$  be distributed independently and identically as elliptically symmetric distribution with density

$$f(\underline{x}) = |\Sigma|^{-1/2} h((\underline{x} - \underline{\mu})' \Sigma^{-1} (\underline{x} - \underline{\mu})). \quad (3.4)$$

Also let  $\lambda_1 \geq \dots \geq \lambda_p$  denote the eigenvalues of  $S_b S_w^{-1}$ . In addition, let  $\theta_1 \geq \dots \geq \theta_p$  denote the eigenvalues of  $\Omega \Sigma^{-1}$  whose multiplicities are given below:

$$\begin{aligned} \theta_1 &= \dots = \theta_{p_1^*} = n\delta_1 \\ \theta_{p_1^*+1} &= \dots = \theta_{p_2^*} = n\delta_2 \\ &\vdots \\ \theta_{p_{t-1}^*+1} &= \dots = \theta_{p_t^*} = n\delta_t \\ \theta_{p_t^*+1} &= \dots = \theta_p = 0 \end{aligned} \quad (3.5)$$

where  $p_j^* = p_1 + \dots + p_j$  ( $j=1,2,\dots,t+1$ ),  $r = p_t^*$ ,  $p = p_1 + \dots + p_{t+1}$  and  $p_0^* = 0$ . In addition, let

$$\begin{aligned} u_{i_h} &= \sqrt{n(2\delta_h^2 + 4\delta_h)}^{-1/2} (\ell_{i_h} - \delta_h) \\ u_{r+j} &= n\ell_{r+j} \end{aligned} \quad (3.6)$$

where  $h = 1,2,\dots,t$ ,  $i_h = p_{h-1}^* + 1, \dots, p_h^*$  and  $j = 1,2,\dots,s-r$  where  $r$  denotes the number of non-zero eigenvalues of  $\Omega \Sigma^{-1}$ . Now, let  $\eta_i = nq_i$  for  $i = 1,2,\dots,k$ . Then, Bai, Krishnaiah and Liang (1984) derived the following expression for the limiting distribution of  $u_1, \dots, u_s$  as  $n \rightarrow \infty$ :

$$f(u_1, \dots, u_s) = \prod_{j=1}^{t+1} \eta_j (u_{p_{j-1}^*+1}, \dots, u_{p_j^*}). \quad (3.7)$$

Here  $\eta_j(\cdot)$ , ( $j=1,2,\dots,t+1$ ), denotes the joint distribution of the eigenvalues of the random matrix  $A_j$ . For  $j=1,2,\dots,t$ , the elements of  $A_j$  are distributed independently as normal with zero means, and variances of the diagonal elements are equal to 1 whereas the variances of the off-diagonal elements are equal to  $1/2$ . In other words, the random matrices  $A_1, \dots, A_t$  are known to be distributed as central Gaussian matrices. Also,  $A_{t+1} : (s-r) \times (s-r)$  is distributed as central Wishart matrix with  $(k-1-r)$  degrees of freedom. Computational aspects of the percentage points of the individual eigenvalues of the central Gaussian matrix and central Wishart matrix are discussed in Krishnaiah (1980). When the underlying distribution is multivariate normal, the expression (3.7) was derived by Hsu (1941). W.Q. Liang (personal communication) found an error in the proof of Hsu. However, Bai (1984) pointed out that the final result of Hsu is correct. Bai, Krishnaiah and Liang (1984) showed that the above result is true even when the observations are distributed independently as elliptically symmetric. From the result of Bai, Krishnaiah and Liang (1984) it is obvious that when  $r$  is the rank of  $\Omega$ ,  $n(\lambda_{r+1} + \dots + \lambda_s)$  is asymptotically distributed as  $\chi^2$  with  $(s-r)(k-1-r)$  degrees of freedom even when the underlying distribution of the observation is elliptically symmetric.

Bai, Krishnaiah and Liang (1984) proposed the following sequential procedure for the rank of  $\Omega$  when  $n_1, \dots, n_k$  tend to infinity such that  $(n_1/n), \dots, (n_k/n)$  tend to (say)  $q_1, \dots, q_k$  respectively. The hypothesis  $\Omega = 0$  is accepted or rejected according as

$$\lambda_1 \leq c_{\alpha_1} \quad (3.8)$$

where

$$P[\lambda_1 \leq c_{\alpha_1} | \Omega = 0] = (1 - \alpha_1). \quad (3.9)$$

If  $\Omega = 0$ , we don't proceed further. If  $\Omega = 0$  is rejected, we accept or reject  $H_1$  according as

$$\lambda_2 \leq c_{\alpha_2} \quad (3.10)$$

where

$$P[\lambda_2 \leq c_{\alpha 2} | H_1; \lambda_1 > c_{\alpha 1}] = (1 - \alpha_2) \quad (3.11)$$

and  $H_t$  denotes the hypothesis that the rank of  $\Omega$  is  $t$ . When  $H_1$  is true, the distribution of  $\lambda_2$  is independent of  $\lambda_1$ . If  $H_1$  is accepted, we don't proceed further. Otherwise, we accept or reject  $H_2$  according as

$$\lambda_3 \leq c_{\alpha 3} \quad (3.12)$$

where

$$P[\lambda_3 \leq c_{\alpha 3} | H_2; \lambda_2 > c_{\alpha 2}] = (1 - \alpha_3). \quad (3.13)$$

We continue this method until a decision is made about the rank of  $\Omega$ .

We now discuss the problem of testing for the rank of the regression matrix. Consider the model

$$Y = XB + E \quad (3.14)$$

where the rows of  $E : n \times p$  are distributed as a multivariate normal with mean vector  $\underline{0}$  and covariance matrix  $\Sigma$ . Also, let  $X : n \times q$  denote the design matrix and  $B : q \times p$  the regression matrix. We assume that  $q \geq p$ . Tintner (1945) derived the LRT statistic for the rank of  $B$  when  $\Sigma$  is known. Anderson (1951) derived the following expression for the LRT statistic to test the hypothesis  $H_r$  which states that the rank of  $B$  is  $r$ :

$$L_1 = \prod_{j=r+1}^p (1 + \lambda_j)^{n/2} \quad (3.15)$$

where  $\lambda_1 \geq \dots \geq \lambda_p$  denote the eigenvalues of  $S_1 S^{-1}$ , and

$$S_1 = Y'X(X'X)^{-1}X'Y \quad (3.16)$$

$$S = Y'[I - X(X'X)^{-1}X']Y. \quad (3.17)$$

Fujikoshi (1977) derived expressions for the asymptotic distributions of the test statistics  $m_1 T_1$ ,  $m_2 T_2$  and  $m_3 T_3$  where

$$\begin{aligned} T_1 &= \sum_{j=r+1}^p \log(1 + \lambda_j) \\ T_2 &= \sum_{j=r+1}^p \lambda_j \\ T_3 &= \sum_{j=r+1}^p \{\lambda_j / (1 + \lambda_j)\}. \end{aligned} \quad (3.18)$$

Here  $m_1, m_2$  and  $m_3$  are certain correction factors. We may choose  $m_1$  to be equal to  $n$ . In deriving the asymptotic distributions, it was assumed that  $\lim_{n \rightarrow \infty} (\Omega/n) = O(1)$  where  $\Omega =$

$B'(X'X)B\Sigma^{-1}$ . The first terms in the asymptotic distributions of  $nT_1, nT_2, nT_3$ , when the null hypothesis is true, are distributed as chi-square distribution with  $(p-r)(q-r)$  degrees of freedom. Fujikoshi also derived nonnull distributions of the above test statistics in terms of normal density and its derivatives when the eigenvalues of  $\Omega$  have multiplicities.

Recently, Krishnaiah, Lin and Wang (1985a) derived the LRT statistics for testing hypothesis on the rank of  $B$  when the underlying distribution is elliptically symmetric; these authors have also investigated the asymptotic distributions of the above statistics. A review of their work is given below.

Let  $E$  be distributed as elliptically symmetric distribution with density

$$f(E) = \frac{1}{|\Sigma|^{n/2}} h(\text{tr} \Sigma^{-1} E'E) \quad (3.19)$$

where  $h(x)$  is strictly decreasing and differentiable function of  $x$ . Also, let

$$\Delta = CB \quad (3.20)$$

where  $C: u \times k$  is known and of rank  $u$ . Let  $H_{1r}$  denote the hypothesis that the rank of  $\Delta$  is  $r$  whereas  $H_{2r}$  denotes the hypothesis that the rows of  $\Delta$  lie in a  $r$ -dimensional plane in  $p$ -dimensional space. Now, let  $\Pi_r(a)$  denote the set of  $n \times p$  matrices of the form  $L = (GF + \underline{a}\underline{b}')D$  where  $|G'G| \neq 0$ ,  $FF' = I_r$ ,  $D: p \times p$  is any positive definite matrix and  $\underline{b}$  is any  $p \times 1$  vector. Then  $H_{1r}$  denotes the hypothesis that  $\Delta \in \Pi_r(0)$  and  $H_{2r}$  denotes the hypothesis that  $\Delta \in \Pi_r(\underline{1})$  where  $\underline{1} = (1, \dots, 1)$ . Now, let

$$M = C(X'X)^{-1}C'$$

$$\hat{B} = (X'X)^{-1}X'Y$$

$$S_h(\hat{B}) = (C\hat{B})'M^{-1}(C\hat{B}) \quad (3.21)$$

$$S_f(\hat{B}) = (C\hat{B})' \{M^{-1} - M^{-1}\underline{1}(\underline{1}'M^{-1}\underline{1})^{-1}\underline{1}'M^{-1}\} C\hat{B}$$

$$S = Y'(I - X(X'X)^{-1}X')Y.$$

Let  $T_4$  and  $T_5$  denote the LRT statistics for testing the hypothesis  $H_{1r}$  against  $H_{1r'}$  for some  $r' > r$  when  $\Sigma$  is known and unknown respectively. Then

$$T_4 = \frac{h(\phi_{r+1} + \dots + \phi_s + \text{tr} \Sigma^{-1} S)}{h(\text{tr} \Sigma^{-1} S)} \quad (3.22)$$

$$T_5 = \prod_{j=r+1}^s (1+d_j)^{-n/2} \quad (3.23)$$

where  $s = \min(u, p)$ ,  $\phi_1 \geq \dots \geq \phi_s$  are the non-zero eigenvalues of  $S_h(\hat{B})\Sigma^{-1}$  and  $d_1 \geq \dots \geq d_s$  are the non-zero eigenvalues of  $S_h(\hat{B})S^{-1}$ . Next, let  $T_6$  and  $T_7$  denote the LRT statistics for testing the hypothesis  $H_{2r}$  against  $H_{2r'}$  for some  $r' > r$  when  $\Sigma$  is known and unknown respectively. Then

$$T_6 = \frac{h(\psi_{r+1} + \dots + \psi_{\bar{s}} + \text{tr} \Sigma^{-1} S)}{h(\text{tr} \Sigma^{-1} S)} \quad (3.24)$$

$$T_7 = \{(1 + \lambda_{r+1}) \dots (1 + \lambda_{\bar{s}})\}^{-n/2} \quad (3.25)$$

where  $\bar{s} = \min(u-1, p)$ ,  $\psi_1 \geq \dots \geq \psi_{\bar{s}}$  are the non-zero eigenvalues of  $S_f(\hat{B})\Sigma^{-1}$  and  $\lambda_1 \geq \dots \geq \lambda_{\bar{s}}$  are the non-zero eigenvalues of  $S_f(\hat{B})S^{-1}$ . Krishnaiah, Lin and Wang (1985a) also derived the LRT statistics analogous to  $T_4, T_5, T_6$  and  $T_7$  when the underlying distribution is complex elliptically symmetric. The above authors also derived asymptotic joint distributions of  $(d_1, \dots, d_s)$  and  $(\lambda_1, \dots, \lambda_{\bar{s}})$ . On the basis of the above results, they pointed out that  $-2\log T_5$  and  $-2\log T_7$  are distributed asymptotically as chi-square. When the underlying distribution is multivariate normal, Rao (1973) derived  $T_6$ .

In a number of situations, it may not be realistic to assume that the joint distribution of the observations  $Y$  is elliptically symmetric. It is more realistic to assume that the rows of  $E$  are distributed independently and identically as elliptically symmetric. The two situations described above become identical when the underlying distribution is multivariate normal. Krishnaiah, Lin and Wang (1985a) have derived asymptotic joint distributions  $(d_1, \dots, d_s)$  and  $(\lambda_1, \dots, \lambda_{\bar{s}})$  when the rows of  $E$  are distributed independently as elliptically symmetric with mean vector  $0$  and the same dispersion matrix.

#### 4 INFERENCE ON THE RANK OF REGRESSION MATRIX USING MODEL SELECTION METHODS

In the model (3.1), we assume that the rows of  $E$  are distributed independently and

identically with mean vector  $\underline{0}$  and covariance matrix  $\Sigma$ . Also, let  $\Delta = CB$  be as defined in the preceding section. In the preceding section, we discussed the problem of testing the hypothesis that the rank of  $\Delta$  is  $r$  where  $r$  is specified. But, situations arise often when the experimenter does not know as to which of the hypotheses  $H_{10}, H_{11}, \dots, H_{1u}$  to test. In these situations, it is of interest to select one of the models  $M_0, M_1, \dots, M_u$  where  $M_j$  denotes the model that the rank of  $\Delta$  is  $j$ . We now give a review of the recent work of Bai, Krishnaiah and Zhao (1986a) for the determination of the rank of  $\Delta$  using model selection methods. Let

$$L(r) = (1/2) \sum_{j=r+1}^s \phi_j + rC_n \quad (4.1)$$

where  $-\frac{1}{2}(\phi_{r+1} + \dots + \phi_s)$  is the logarithm of the LRT statistic for testing the hypothesis that the rank of  $\Delta$  is  $r$  when  $\Sigma$  is known and the underlying distribution is multivariate normal. The statistics are as defined in the preceding section. Also,  $C_n$  satisfies the following conditions:

$$(i) \lim_{n \rightarrow \infty} \{C_n / \log n\} = \infty$$

$$(ii) \lim_{n \rightarrow \infty} \{C_n / \lambda_*\} = 0 \quad (4.2)$$

$$(iii) \lim_{n \rightarrow \infty} \{\lambda_* / \log n\} = \infty$$

where  $\lambda_*$  denotes the smallest eigenvalue of  $X'X$ . Then, according to the procedure of Bai, Krishnaiah and Zhao (1986a), the rank of  $\Delta$  when  $\Sigma$  is known is estimated with  $\hat{q}$  where

$$L(\hat{q}) = \min\{L(0), L(1), \dots, L(s)\}. \quad (4.3)$$

The above authors also showed that  $\hat{q}$  defined above is a consistent estimate of the rank of  $\Delta$ .

When  $\Sigma$  is unknown, let

$$L^*(r) = \frac{n}{2} \sum_{j=r+1}^s \log(1+d_j) + rC_n \quad (4.4)$$

where  $d_1 \geq \dots \geq d_s$  are the first  $s$  largest eigenvalues of  $S_h(B)S^{-1}$  defined in the preceding section and  $C_n$  satisfies the following conditions:

$$(i) \lim_{n \rightarrow \infty} (C_n / \log n) = \infty$$

$$(ii) \lim_{n \rightarrow \infty} (C_n) < (n/3) \log 2 \quad (4.5)$$

$$(iii) \lim_{n \rightarrow \infty} (C_n / \lambda_n^*) = 0.$$

We also make the following assumptions on  $\lambda^*$  (largest eigenvalue of  $X'X$ ) and  $\lambda_n^*$ :

$$(i) \lim_{n \rightarrow \infty} (\lambda_n^* / \log n) = \infty$$

(4.6)

$$(ii) \lambda^* = O(n \log n / \log \log n)$$

Then, Bai, Krishnaiah and Zhao (1986a) proposed using  $\hat{q}$  as an estimate of the rank of  $B$  where

$$L^*(\hat{q}) = \min \{L^*(0), \dots, L^*(s)\}.$$

They also proved that  $\hat{q}$  is a consistent estimate of the rank of  $B$ .

We may consider alternative model selection criteria similar to those considered by Akaike (1972), Rissanen (1978) and Schwartz (1978) in some other problems.

Next consider the case when  $X$  is also stochastic and the rows of  $(Y \ X)$  are distributed independently as multivariate normal with mean vector  $0$  and unknown covariance matrix. When  $B$  is not of full rank, Izenman (1974) considered the problem of estimation of  $B$  and asymptotic distribution of the estimate of  $B$ . We can propose model selection procedures, similar to those discussed in the present section, to determine the rank of  $B$ .

## 5. REDUCTION OF DIMENSIONALITY UNDER FANOVA MODEL

Consider the following two-way classification model with one observation per cell:

$$x_{ij} = \mu + \alpha_i + \beta_j + \eta_{ij} + \varepsilon_{ij} \quad (5.1)$$

for  $i = 1, 2, \dots, r$ ,  $j = 1, 2, \dots, s$ , where

$$\sum_{i=1}^r \alpha_i = \sum_{j=1}^s \beta_j = \sum_{i=1}^r \eta_{ij} = \sum_{j=1}^s \eta_{ij} = 0. \quad (5.2)$$

Here  $\mu, \alpha_i, \beta_j$  and  $\eta_{ij}$  respectively denote the general mean, effect due to  $i$ -th row, effect due to  $j$ -th column and interaction in  $i$ -th row and  $j$ -th column respectively. Without loss of generality, we assume that  $r \leq s$ . The problem of finding the rank of the interaction matrix  $\eta = (\eta_{ij})$  is of interest and received attention in the literature. The usual F test statistics to test the hypotheses of no row effect and no column effect were proposed in the literature under the assumption of no interactions. If there is interaction, then the F statistics are no longer distributed as central F distributions even when the null hypotheses are true and so the usual tests are no longer valid. So, it is of interest to test the hypothesis that the rank of  $\eta$  is zero; this problem is known in the literature as testing for additivity. Fisher and MacKenzie (1923), Tukey (1949) and Williams (1952) are the early workers on the problem of testing for additivity when  $\eta$  has special structures. When  $\eta \neq 0$ , knowledge of the rank of  $\eta$  will help to estimate the parameters more efficiently. So, it is of interest to test for the rank of  $\eta$ . We will now discuss this problem.

Suppose  $\eta$  is of rank  $c$ . Then it is known, by singular value decomposition of the matrix, that

$$\eta = \theta_1 \underline{\mu}_1 \underline{\nu}_1' + \dots + \theta_c \underline{\mu}_c \underline{\nu}_c' \quad (5.3)$$

where  $\theta_1^2 \geq \dots \geq \theta_c^2$  are the eigenvalues of  $\eta\eta'$ ,  $\underline{\mu}_j$  and  $\underline{\nu}_j$  are the eigenvectors of  $\eta\eta'$  and  $\eta'\eta$  corresponding to  $\theta_j^2$ . Now, let  $\lambda_1 \geq \dots \geq \lambda_{r-1}$  denote the non-zero eigenvalues of  $DD'$  where  $D = (d_{ij})$  and  $d_{ij} = x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..}$ . Gollob (1968) considered the problem

of testing the hypotheses  $\theta_j = 0$  and his tests are based upon the assumption that  $\ell_j$ 's are distributed independently as chi-square variables. But the above assumption is not correct. Mandel (1969) proposed heuristically to examine the magnitude of  $\ell_j / \gamma_j \hat{\sigma}^2$  to test for  $\theta_j = 0$  where  $\gamma_j = E(\ell_j)$  and  $\hat{\sigma}^2 = (\ell_{c+1} + \dots + \ell_{r-1}) / (\gamma_{c+1} + \dots + \gamma_{r-1})$ . But the distributions of the above test statistics are not only complicated but also involve nuisance parameters. Corsten and van Eijnsbergen (1972) derived the following likelihood ratio test. Accept or reject  $H: \theta_1 = \dots = \theta_c = 0$  according as

$$L_1 \leq c_{1\alpha} \quad (5.4)$$

where

$$P[L_1 \leq c_{1\alpha} | H] = (1-\alpha) \quad (5.5)$$

and  $L_1 = (\ell_1 + \dots + \ell_c) / (\ell_1 + \dots + \ell_{r-1})$ . When  $c = 1$ , the likelihood ratio test statistic was derived independently by Johnson and Graybill (1972). Yochmowitz and Cornell (1978) discussed the likelihood ratio test statistic for testing the hypothesis  $\theta_j = 0$  against the alternative  $\theta_j \neq 0$  and  $\theta_{j+1} = \dots = \theta_c = 0$ . When  $H$  is true, it is known (e.g., see Johnson and Graybill (1973)) that  $\ell_1, \dots, \ell_{r-1}$  are jointly distributed as the joint distribution of the eigenvalues of the central  $(r-1) \times (r-1)$  Wishart matrix  $W$  with  $(s-1)$  degrees of freedom and  $E(W) = (s-1)I_{r-1}$ . Schuurmann, Krishnaiah and Chattopadhyay (1973) derived the exact distribution of  $\ell_1 / (\ell_1 + \dots + \ell_{r-1})$  and  $\ell_{r-1} / (\ell_1 + \dots + \ell_{r-1})$  when  $H$  is true and computed some of the percentage points of the above statistic. Krishnaiah and Schuurmann (1974) derived the exact distributions of  $\ell_j / (\ell_1 + \dots + \ell_{c-1})$  for  $j = 2, 3, \dots, c-1$  when  $H$  is true. Schuurmann, Krishnaiah and Chattopadhyay (1973) proposed the following simultaneous test procedure in the spirit of the simultaneous test procedures of Krishnaiah and Waikar (1971a,b) in the area of principal component analysis. Accept or reject  $\theta_i = 0$  according as

$$\frac{\ell_i}{\ell_1 + \dots + \ell_{c-1}} \leq c_{2\alpha} \quad (5.6)$$

where

$$P \left\{ \frac{\ell_1}{\ell_1 + \dots + \ell_{c-1}} \leq c_{2\alpha} \mid H \right\} = (1-\alpha). \quad (5.7)$$

For details of other simultaneous test procedures, the reader is referred to Krishnaiah and Yochmowitz (1980).

We will now review the recent work of Rao (1985) on a more general problem of reduction of dimensionality.

Let  $Y : n \times p$  be a random matrix which is distributed as multivariate normal with  $E(Y) = M$  and the covariance matrix of  $\underline{y}$  is  $C_y \Sigma$  where  $\underline{y}$  is the vector obtained by writing the rows of  $Y$  vertically one below the other starting from the first and  $C$  is a known positive definite matrix. Also, let  $S : p \times p$  be distributed independent of  $Y$  as central Wishart matrix with  $s$  degrees of freedom and  $E(S) = s\Sigma$ . Under the above model, Rao (1985) derived the likelihood ratio tests for testing the hypothesis  $H$  where

$$H : M = X\psi + \phi W' + \Gamma \quad (5.8)$$

where  $\Sigma$  has general structure and has the structure of the form

$$\Sigma = \sigma_1^2 V_1 V_1' + \dots + \sigma_f^2 V_f V_f' \quad (5.9)$$

where  $\sigma_1^2, \dots, \sigma_f^2$  are unknown and  $V_i : p \times g_i$  ( $i = 1, 2, \dots, f$ ) is known matrix of rank  $g_i$  such that  $p = g_1 + \dots + g_f$ .

In (5.8),  $X : n \times b$  is a known matrix of rank  $b$ ,  $W : p \times c$  is a given matrix of rank  $c$ ,  $\psi$  and  $\phi$  are matrices of unknown parameters, and  $\Gamma$  is a matrix of specified rank  $r \leq \min(k-b, p-c)$ . If  $X$  is a  $n \times 1$  vector of unities and  $W$  is null matrix, the above problem reduces to the problem of specifying the dimensionality of row mean vectors in  $M$  considered by Fisher (1939), Fujikoshi (1974), Krishnaiah,

Lin and Wang (1985a) and others. If  $X$  is a  $n \times 1$  vector of unities and  $W$  is a  $p \times 1$  vector of unities, then  $H$  is the hypothesis specifying the rank of interaction in two-way classification with one observation per cell and this problem was considered when  $\Sigma = \sigma^2 I$  and  $C = I$ .

## 6. RANK OF COVARIANCE MATRIX OF RANDOM EFFECTS IN ONE-WAY COMPONENTS OF COVARIANCE MODEL

Consider the one-way components of covariance model

$$x_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad (6.1)$$

for  $i = 1, 2, \dots, k$ ,  $j = 1, 2, \dots, m_i$  where  $\mu$  is the general mean vector,  $\alpha_i$  is the vector of random effects, and  $\varepsilon_{ij}$  is vector of errors, and  $x_{ij}$  denotes  $j$ -th observation in  $i$ -th group. Also,  $\alpha_i$  and  $\varepsilon_{ij}$  are distributed independent of each other as multivariate normal with  $E(\alpha_i) = E(\varepsilon_{ij}) = 0$  and covariance matrices given by  $\psi = E(\alpha_i \alpha_i')$  and  $E(\varepsilon_{ij} \varepsilon_{ij}') = \Sigma_1$ . We also assume that  $E(\alpha_i \alpha_{i'}) = 0$  for  $i \neq i'$  and  $E(\varepsilon_{ij} \varepsilon_{i'j'}) = 0$  for  $i \neq i'$  and/or  $j \neq j'$ . The covariance matrix of  $x_{ij}$  is given by  $\Sigma_2$  where

$$\Sigma_2 = \psi + \Sigma_1 \quad (6.2)$$

We assume that  $\psi$  is not of full rank and we are interested in finding out the rank of  $\psi$ . If the rank of  $\psi$  is  $r$ , then there exists a full rank matrix  $B: (p-r) \times p$  such that  $B\psi = 0$ . If the rank of  $\psi$  is zero, then we conclude that there is no difference between the effects of the groups. Knowledge about the rank of  $\psi$  will help to estimate  $\psi$  more efficiently.

When  $m_1 = m_2 = \dots = m_k$ , the between groups sums of squares and cross products (SP) matrix and within group SP matrix are given by  $S_b$  and  $S_w$  respectively where

$$S_b = m \sum_{i=1}^k (\bar{x}_{\cdot i} - \bar{x})(\bar{x}_{\cdot i} - \bar{x})'$$

$$S_w = \sum_{i=1}^k \sum_{j=1}^m (\bar{x}_{ij} - \bar{x})(\bar{x}_{ij} - \bar{x})' \quad (6.3)$$

$$m\bar{x}_{\cdot i} = \sum_{j=1}^m x_{ij}, \quad km\bar{x}_{\cdot i} = \sum_{i=1}^k \sum_{j=1}^m x_{ij}$$

Then,  $S_b$  and  $S_w$  are distributed independently as central Wishart matrices with  $(k-1)$  and  $k(m-1)$  degrees of freedom respectively,  $E(S_b/(k-1)) = \Sigma_2$ , and  $E(S_w/k(m-1)) = \Sigma_1$  and  $\Sigma_2 = \Sigma_1 + m\psi$ . When the sample sizes are unequal,  $S_b$  is not distributed as Wishart matrix. When  $m$ 's are equal, Anderson ((1984),(1985)) has derived the likelihood ratio test statistic for testing the hypothesis that the rank of  $\psi$  is not greater than  $r$ . Schott and Saw (1984) derived the likelihood ratio test for rank  $(\psi \leq r)$  against the alternative rank  $(\psi) = r+1$ .

We now discuss a more general problem considered by Rao (1983) and Zhao, Krishnaiah and Bai (1985b). Let  $S_1$  and  $S_2$  be distributed independently as central Wishart matrices with  $n_1$  and  $n_2$  degrees of freedom respectively and let  $E(S_i/n_i) = \Sigma_i$  for  $i = 1, 2$ . Also, let  $\Sigma_2 = \Gamma + \Sigma_1$  where  $\Gamma$  is a nonnegative definite matrix. Then, we are interested in finding the rank of  $\Gamma$ . Rao (1983) proposed a modified LRT statistic for testing the hypothesis that the rank of  $\Gamma$  is a specified value. We will now discuss the model selection method proposed by Zhao, Krishnaiah and Bai (1985b) for estimating the rank of  $\Gamma$ . Let  $\delta_1 \geq \dots \geq \delta_p$  denote the eigenvalues of  $S_1 S_2^{-1} n_2/n_1$ . Also, let

$$L_q = \prod_{i=1+\min(q, \tau)}^p \{(\alpha_n + \beta_n \delta_i)^{-n/2} \delta_i^{-n/2}\} \quad (6.4)$$

where  $\tau$  denotes the number of  $\delta_i$ 's which are greater than one,  $\alpha_n = n_1/n$ ,  $\beta_n = n_2/n$  and  $n = n_1 + n_2$ . In addition, let

$$L_{qt} = \prod_{i=1+\min(q,T)}^{1+\min(t,T)} \{(\alpha_n + \beta_n \delta_i)^{-n/2} e^{-n\delta_i^2/2}\} \quad (6.5)$$

Zhao Krishnaiah and Bai (1985b) showed that  $L_q$  is the likelihood ratio test statistic for testing  $H_q$  against the alternative that  $\Sigma_1$  and  $\Sigma_2$  are arbitrary and  $L_{qt}$  is the likelihood ratio test statistic for testing  $H_q$  against  $H_t(q < t)$  where  $H_t$  denotes the hypothesis that the rank of  $\Gamma$  is equal to  $t$ . Now let.

$$EDC(a, C_n) = -\log L_a + v(a, p)C_n \quad (6.6)$$

where  $v(a, p) = (1/2)a(2p-a+1)$  and  $C_n$  satisfies the following conditions

$$(i) \lim_{n \rightarrow \infty} (C_n/n) = 0$$

(6.7)

$$(ii) \lim_{n \rightarrow \infty} (C_n/\log \log n) = \infty$$

Zhao, Krishnaiah and Bai (1985b) estimated the unknown rank of  $\Gamma$  with  $\hat{q}$  where

$$EDC(\hat{q}, C_n) = \min \{EDC(0, C_n), \dots, EDC(p-1, C_n)\} \quad (6.8)$$

They have also proved that  $\hat{q}$  is strongly consistent. The above procedure can be used to draw inference on the rank of the covariance matrix of the vector of random effects in one way components of covariance model.

## 7. SELECTION OF ORIGINAL VARIABLES UNDER MULTIVARIATE REGRESSION MODEL

In the area of univariate regression analysis, it is of interest to select variables which are important for prediction. Reviews of the literature on some methods of selection of variables are given in Krishnaiah (1982) and Thompson ((1978a),(1978b)). In this section, we review procedures for selection of independent variables which are important for prediction of a set of dependent variables under classical multivariate regression model.

Consider the multivariate regression model (3.14) where  $X = [x_1, \dots, x_q]$  and  $x_i : n \times 1$  is vector of  $n$  independent observations on the  $i$ -th independent variable  $x_i$ . Also, let  $Y = [y_1, \dots, y_p]$  where  $y_i : n \times 1$  denotes the vector of  $n$  independent observations on  $i$ -th dependent variable. Then, it is of interest to find out as to which of the variables  $x_1, \dots, x_q$  are important. We can use Roy's largest root test,  $T_{\max}^2$  test or Krishnaiah's finite intersection tests for the selection of important variables. Now, let  $B' = (\beta_1, \dots, \beta_q)$  where  $\beta_i$  is of order  $p \times 1$ . Also, let  $H_1 : \beta_1 = 0$  and

$$T_1^2 = \frac{(n-q) \hat{\beta}_1' S^{-1} \hat{\beta}_1}{e_{11}} \quad (7.1)$$

where  $\hat{B} = (\hat{\beta}_1, \dots, \hat{\beta}_q)' = (X'X)^{-1}(X'Y)$ ,  $S = (S_{ij}) = Y'(I - X(X'X)^{-1}X')Y$  and  $e_{11}$  is the covariance matrix of  $\hat{\beta}_1$ . According to Roy's largest root test, we accept or reject  $H_1$  according as

$$T_1^2 \begin{matrix} < \\ > \end{matrix} c_\alpha \quad (7.2)$$

where  $c_\alpha$  is chosen such that

$$P[(n-q)C_L(S_1, S^{-1}) \leq qc_\alpha | H] = (1-\alpha) \quad (7.3)$$

where  $H = \bigcap_{i=1}^q H_i$  and  $C_L(A)$  denotes the largest eigenvalue of  $A$  and  $S_1 = Y'X(X'X)^{-1}X'Y = \hat{B}'(X'X)\hat{B}$ . Percentage points of  $c_\alpha$  are given in Krishnaiah (1980). If we use  $T_{\max}^2$  test (e.g., see Krishnaiah (1969) and Siotani (1959)), we accept or reject  $H_1$  according as

$$T_i^2 \leq c_{\alpha 1}$$

where

$$P[T_i^2 \leq c_{\alpha 1}; i = 1, 2, \dots, q | H] = (1 - \alpha).$$

Approximate values of  $c_{\alpha 1}$  can be obtained from the results of Sictani ((1959),(1960),(1961)) for some cases. We conclude that the independent variable  $x_i$  is important or unimportant for prediction of  $(y_1, \dots, y_p)$  according as  $H_i$  is rejected or accepted. We now discuss Krishnaiah's finite intersection tests (Krishnaiah (1965)) for the selection of variables. For an illustration of the application of the finite intersection test, the reader is referred to Schmidhammer (1982).

Let  $\Sigma_k$  denote the top  $k \times k$  left-hand corner of  $\Sigma = (\sigma_{ij})$  and  $\sigma_{k+1}^2 = |\Sigma_{k+1}| / |\Sigma_k|$  for  $k = 0, 1, \dots, p-1$  with  $|\Sigma_0| = 1$ . Also, let  $Y_j = [y_1, \dots, y_j]$ ,  $X_j = [x_1, \dots, x_j]$ , and  $B_j = [\beta_1, \dots, \beta_j]$  for  $j = 1, 2, \dots, p$ . In addition, let

$$\zeta_j = \Sigma_j^{-1} \begin{pmatrix} \sigma_{1,j+1} \\ \vdots \\ \sigma_{j,j+1} \end{pmatrix} \quad (7.6)$$

for  $j = 1, 2, \dots, p-1$ ,  $\zeta_0 = 0$ . We know, that the conditional distribution of  $y_{j+1}$ , given  $Y_j$ , is distributed as multivariate normal with covariance matrix  $\sigma_{j+1}^2 I_n$  and the mean vector

$$E_c(y_{j+1}) = X_{j+1} \eta_{j+1} + Y_j \zeta_j = [X, Y_j] \begin{pmatrix} \eta_{j+1} \\ \zeta_j \end{pmatrix} \quad (7.7)$$

where  $\eta_{j+1} = \beta_{j+1} - B_j \zeta_j$  with the understanding that  $\eta_1 = \beta_1$ . Now, let  $H_{ij} : c'_{ij} \eta_j = 0$  where  $c'_{ij} = (c_{i1}, \dots, c_{iq})$  for  $i = 1, 2, \dots, q$  with

$$c_{ih} = \begin{cases} 0 & h \neq i \\ 1 & h = i. \end{cases}$$

Then, the hypothesis  $H_1$  can be expressed as  $H_1 = \bigcap_{j=1}^p H_{1j}$ . So, the problem of testing the hypotheses  $H_1, \dots, H_q$  simultaneously is equivalent to testing the hypotheses  $H_{ij}$  simultaneously. Now, let

$$F_{ij} = \frac{(\hat{c'_{ij} \eta_j})^2 (n-j-q+1)}{d_{ij} s_j^2} \quad (7.8)$$

where  $d_{ij} \sigma_j^2$  is the variance of  $\hat{c'_{ij} \eta_j}$ ,  $\hat{\eta_j}$  is the least square estimate of  $\eta_j$  under the model (7.7), and  $s_{j+1}^2 = |S_{j+1}| / |S_j|$  where  $S_j$  is the top  $|j|$  left-hand corner of  $S$ . Then, we accept or reject  $H_{ij}$  according as

$$F_{ij} \leq F_\alpha \quad (7.9)$$

where

$$\begin{aligned} P[F_{ij} \leq F_\alpha, i = 1, 2, \dots, q, j = 1, 2, \dots, p | H] \\ = \prod_{j=1}^p P[F_{ij} \leq F_\alpha, i = 1, 2, \dots, q | H] \\ = (1-\alpha). \end{aligned} \quad (7.10)$$

When  $H$  is true, the joint distribution of  $F_{1j}, \dots, F_{qj}$  is a multivariate  $F$  distribution with  $(1, n-q-j+1)$  degrees of freedom. Evaluation of the probability integrals of the multivariate  $F$  distribution was discussed in Krishnaiah and Armitage (1970). The hypothesis  $H_1$  is accepted if  $H_{11}, \dots, H_{1p}$  are accepted and it is rejected otherwise. If  $H_1$  is rejected, then we conclude that the independent variable  $x_1$  is important for prediction of the set  $(y_1, \dots, y_p)$  of dependent

variables. One may use the step-down procedure proposed by J. Roy (1958) also but the lengths of the confidence intervals associated with the finite intersection tests are shorter than the lengths of the corresponding confidence intervals associated with the step-down procedure. Fujikoshi (1985) proposed a procedure, based on an information theoretic criterion, to select a subset of variables which are important for discrimination. Rao (1948) proposed a procedure to find out as to whether the addition of some independent variables makes a significant contribution in prediction of dependent variables.

#### 8. COMMENTS ON STEPWISE PROCEDURES FOR SELECTION OF VARIABLES IN DISCRIMINANT ANALYSIS

In this section, we discuss the stepwise procedures for the selection of variables in the area of discriminant analysis for several groups. These procedures are used widely since computer programs for the implementation of these procedures are available in the BMD and SPSS packages. Stepwise procedures for the selection of variables in discriminant analysis were proposed in the literature in a similar way as the corresponding procedures in the regression analysis (Krishnaiah(1982)). We will discuss a stepwise procedure below.

Consider the following model:

$$E(\underline{y}_j) = A \underline{\theta}_j \quad (8.1)$$

where

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_k \end{pmatrix} \quad (8.2)$$

In the matrix  $A_j$ ,  $n_j \times k$  the elements in  $i$ -th column are equal to one and other elements in the matrix are zero. Also,  $\underline{\theta}_j = (\underline{\mu}_{1j}, \dots, \underline{\mu}_{kj})$ ,  $\underline{y}'_j = (x_{1j1}, \dots, x_{1jn_1}, \dots, x_{kj1}, \dots, x_{kjn_k})$  and  $x_{ijt}$  denotes observation on  $j$ -th variable,  $t$ -th individual and  $i$ -th group. Let  $H_j : C \underline{\theta}_j = \underline{0}$  where

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix} : (k-1) \times k \quad (8.3)$$

Let  $F_j$  denote the usual F statistic used for testing the hypothesis  $H_j$ . Then

$$F_j = \frac{b_{jj}(n-k)}{w_{jj}(k-1)} \quad (8.4)$$

where  $W = (w_{ij})$  and  $B = (b_{ij})$  are the within group SP matrix and between group SP matrix respectively. The likelihood ratio statistic for testing  $H_j$  is given by  $\Lambda(x_j)$  where

$$\Lambda(x_j) = \frac{w_{jj}}{t_{jj}} \quad (8.5)$$

and  $t_{jj} = b_{jj} + w_{jj}$ . Obviously,

$$F_j = \frac{(1 - \Lambda(x_j))(n-k)}{(k-1)} \quad (8.6)$$

If  $\max(F_1, \dots, F_p) \leq F_{1\alpha}$ , we declare that none of the variables are important for discrimination and we don't proceed further. Otherwise, we select the variable corresponding to the maximum of  $F_1, \dots, F_p$  as the most important. For example, let this variable be  $x_1$ . At the second stage we test to find out as to whether any of the remaining variables  $x_2, x_3, \dots, x_p$  give additional information for discrimination between the populations. A measure of the degree of additional information is provided by

$$\Lambda_{1 \cdot 1} = \frac{\Lambda(x_1, x_j)}{\Lambda(x_1)} \quad (8.7)$$

where

$$\Lambda(x_1, x_j) = \frac{\begin{vmatrix} w_{11} & w_{1j} \\ w_{j1} & w_{jj} \end{vmatrix}}{\begin{vmatrix} t_{11} & t_{1j} \\ t_{j1} & t_{jj} \end{vmatrix}} \quad (8.8)$$

In (8.7),  $\Lambda(x_1, x_j)$  is the likelihood ratio test statistic for testing the hypothesis that the mean vectors of  $(x_1, x_j)$  are the same in all populations. It can be viewed as a measure of the discriminating ability of  $x_1$  and  $x_j$  whereas  $\Lambda(x_1)$  is a measure of the degree of discrimination of the variable  $x_1$ . As the value of  $\Lambda(x_1, x_j)$  decreases, the discriminating ability of  $x_1$  and  $x_j$  increase. We can write  $\Lambda_{j \cdot 1}$  as

$$\Lambda_{j \cdot 1} = \frac{w_{j \cdot 1}}{t_{j \cdot 1}} \quad (8.9)$$

where  $w_{j \cdot 1} = w_{jj} - w_{j1} w_{11}^{-1} w_{1j}$  and  $t_{j \cdot 1} = t_{jj} - t_{j1} t_{11}^{-1} t_{1j}$ . Now let,

$$F_{j \cdot 1} = \frac{b_{j \cdot 1} (n-k-1)}{w_{j \cdot 1} (k-1)} \quad (8.10)$$

where  $b_{j \cdot 1} = t_{j \cdot 1} - w_{j \cdot 1}$  is the adjusted between group sum of squares. We can write (8.10) as

$$F_{j \cdot 1} = \frac{(n-k-1)}{(k-1)} \frac{1 - \Lambda_{j \cdot 1}}{\Lambda_{j \cdot 1}} \quad (8.11)$$

The above statistic (see Rao (1973)) is nothing but the statistic used to test the hypothesis

$$H_{j \cdot 1} : \mu_{\cdot 1} - \beta_{j \cdot 1} \mu_{\cdot j} = \mu_{\cdot 1} - \beta_{j \cdot 1} \mu_{\cdot j} \quad (8.12)$$

where  $\beta_{j,1} = \sigma_{j1} \sigma_{11}^{-1}$ . If  $\max(F_{2,1}, \dots, F_{p,1}) \leq F_{2\alpha}$  we declare that none of the variables  $x_2, x_3, \dots, x_p$  are important; here  $F_{2\alpha}$  is the upper  $\alpha\%$  point of the central F distribution with  $(k-1, n-k-1)$  degrees of freedom. If  $\max(F_{2,1}, \dots, F_{p,1}) > F_{2\alpha}$  the variable corresponding to the maximum of  $F_{2,1}, \dots, F_{p,1}$  is declared to be important. For simplicity of notation, let us assume that this variable is, say,  $x_2$ . After having selected  $x_2$ , we will test whether the variable (in this case  $x_1$ ) selected at the first stage is good for discrimination in presence of the variable  $x_2$ ; this is the third step. This can be tested by using the following test statistic:

$$F_{1,2} = \frac{b_{1,2}(n-k-1)}{w_{1,2}(k-1)} \quad (8.13)$$

where  $t_{1,2} = t_{11} - t_{12} t_{22}^{-1} t_{21}$ ,  $w_{1,2} = w_{11} - w_{12} w_{22}^{-1} w_{21}$  and  $b_{1,2} = t_{1,2} - w_{1,2}$ . We decide to retain or exclude  $x_1$  from the selected subset according as

$$F_{1,2} \geq F_{2\alpha}^* \quad (8.14)$$

Here we note that

$$F_{1,2} = \frac{(n-k-1)}{(k-1)} \frac{(1-\Lambda_{1,2})}{\Lambda_{1,2}} \quad (8.15)$$

where

$$\Lambda_{1,2} = \frac{w_{1,2}}{t_{1,2}} \quad (8.16)$$

$w_{1,2} = w_{11} - w_{12} w_{22}^{-1} w_{21}$  and  $t_{1,2} = t_{11} - t_{12} t_{22}^{-1} t_{21}$ . If  $\Lambda_{1,2}^* = 1/\Lambda_{1,2}$ , then,

$$F_{1,2} = \frac{(n-k-1)(\Lambda_{1,2}^* - 1)}{(k-1)} \quad (8.17)$$

In the fourth step, we either select one of the variables  $x_3, \dots, x_p$  or decide not to select any more on the basis of the discriminating ability of these variables individually in presence of  $x_1$  and  $x_2$ . If we discard  $x_1$  at the third step, then we consider the discriminating ability of the variables  $x_3, \dots, x_p$  in presence of  $x_2$  only. This procedure is continued until a decision is made not to select any more variables or all the variables are selected. Suppose, after a few stages, we selected  $x_3, x_4, \dots, x_j$  and  $x_j$  is the latest addition to the selected subset. Then, we test whether  $x_3, x_4, \dots, x_{j-1}$  are individually important in presence of the remaining variables. For example, we test whether  $x_4$  is important in presence of the variables  $x_3, x_5, x_6, \dots, x_j$ . The statistic used to test whether  $x_i$  ( $i = 3, 4, \dots, j-1$ ) is important is given by

$$F_{i \cdot (3, 4, \dots, j)} = \frac{b_{i \cdot (3, 4, \dots, j)} (n-k-j+3)}{w_{i \cdot (3, 4, \dots, j)} (k-1)} \quad (8.18)$$

with the understanding that the suffix  $i$  does not occur in the set  $(3, 4, \dots, j)$ . Let

$$\Lambda_{i \cdot (3, 4, \dots, j)} = \frac{\Lambda(x_3, x_4, \dots, x_i, \dots, x_j)}{\Lambda(x_3, x_4, \dots, x_{i-1}, x_{i+1}, \dots, x_j)} \quad (8.19)$$

where  $\Lambda(x_3, x_4, \dots, x_i, \dots, x_j)$  is the ratio of the determinant of the within group SP matrix based upon the variables  $x_3, x_4, \dots, x_i, \dots, x_p$  and the determinant of the total SP matrix based upon the same variables. Similarly  $\Lambda(x_3, x_4, \dots, x_{i-1}, x_{i+1}, \dots, x_j)$  can be defined. So,

$$\Lambda_{i \cdot (3, 4, \dots, j)} = \frac{w_{i \cdot (3, 4, \dots, j)}}{t_{i \cdot (3, 4, \dots, j)}} \quad (8.20)$$

Hence

$$F_{i \cdot (3, 4, \dots, j)} = \frac{(n-k-j+3)}{(k-1)} \frac{(1 - \Lambda_{i \cdot (3, 4, \dots, j)})}{\Lambda_{i \cdot (3, 4, \dots, j)}} \quad (8.21)$$

The variable  $x_i$  is retained or excluded according as  $F_{i(3,4,\dots,j)}$  is greater than or less than upper  $100\alpha\%$  point of the central  $F$  distribution with  $(k-1, n-k-j+3)$  degrees of freedom. At any stage, we can test whether all the variables selected together will discriminate between the groups by using many standard procedures. For example, suppose  $x_1, x_2, \dots, x_q$  are selected. Then, we compute  $B_{11}$  and  $W_{11}$  which are respectively between group SP matrix and within group SP matrix based on  $x_1, \dots, x_q$ . They are given by

$$B_{11} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & b_{22} & \dots & b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{q1} & b_{q2} & \dots & b_{qq} \end{bmatrix}, \quad W_{11} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1q} \\ w_{21} & w_{22} & \dots & w_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ w_{q1} & w_{q2} & \dots & w_{qq} \end{bmatrix} \quad (8.22)$$

We can test whether the variables  $x_1, \dots, x_q$  together will discriminate between the populations by computing various functions of the eigenvalues of  $B_{11}W_{11}^{-1}$ . Some of these functions are  $C_L(B_{11}W_{11}^{-1})$ ,  $\text{tr}(B_{11}W_{11}^{-1})$ ,  $\text{tr}(B_{11}(B_{11}+W_{11})^{-1})$  and  $|B_{11}(B_{11}+W_{11})^{-1}|$ . One can also use finite intersection tests.

We now have a critical look at the stepwise procedure for the selection of variables. At the first stage of the procedure, we choose the critical value  $F_{1\alpha}$  such that

$$P[F_j \leq F_{1\alpha} | H_j] = (1-\alpha). \quad (8.23)$$

Here, the hypotheses  $H_1, \dots, H_p$  are tested individually. Since the decision not to select or select any variable at the first stage is based upon whether or not all the hypotheses are accepted simultaneously, it would be a natural thing to test them simultaneously and choose the critical value  $F_{1\alpha}$  such that

$$P[F_j \leq F_{1\alpha} : j = 1, 2, \dots, p | \bigcap_{j=1}^p H_j] = (1-\alpha). \quad (8.24)$$

The joint distribution of  $F_1, \dots, F_p$  is not only complicated but also involves nuisance parameters. But, we can use Bonferroni's inequality to compute an upper bound on  $F_{1\alpha}$ . At the first stage, we select one variable only as the most important and no decision is made about other variables. But, this "most important variable" may be discarded at a later stage. So, there is some inconsistency in this method and we will discuss this point later. At the second stage, the critical value  $F_{2\alpha}$  is chosen such that

$$P[F_{j,1} \leq F_{2\alpha} | H_{j,1}] = (1-\alpha). \quad (8.25)$$

We go to the second stage if and only if  $\max(F_1, \dots, F_p) \geq F_{1\alpha}$ . So, at the second stage, we have to compute the following conditional probabilities instead of (8.23) even if we are testing the hypotheses  $H_{j,1}$  individually:

$$P[F_{j,1} \leq F_{2\alpha} | \max(F_1, \dots, F_p) \geq F_{1\alpha}]. \quad (8.26)$$

It is quite complicated to compute the above probabilities. Apart from it, we have to test  $H_{2,1}, \dots, H_{p,1}$  simultaneously instead of testing them individually. At the second stage, we select the variable (say  $x_2$ ) corresponding to  $\max(F_{2,1}, \dots, F_{p,1})$ . The statistic  $F_{j,1}$  for any given  $j$  is useful for testing whether the variable  $x_j$  gives additional information for discrimination between the groups in presence of the important variable  $x_1$ . But, the variable  $x_1$  which is declared to be the most important at the first stage may be discarded as being unimportant at a later stage and so the procedure may not be meaningful. Apart from it, the choice of the critical values is very arbitrary and we cannot say what the Type I error of this procedure is. In view of the points raised above, we do not recommend the use of the above stepwise procedures. Krishnaiah (1982) discussed the disadvantages of using forward selection and backward selection procedures for selection of variables under univariate regression models. Similar criticism applies for forward selection and backward selection procedures for selection of variables in discriminant analysis.

## 9. TESTS FOR THE RANK OF THE CANONICAL CORRELATION MATRIX

It is known that multiple correlation coefficient is the maximum correlation between a variable and linear combinations of a set of variables. Hotelling ((1935),(1936)) generalized the above concept to two sets of variables  $x'_1:1xp_1$  and  $x'_2:1xp_2$  and introduced canonical correlation analysis. Canonical correlation analysis is useful in studying the relationship between the two sets of variables. Let the covariance matrix of  $\underline{x} = (\underline{x}_1, \underline{x}_2)$  be  $\Sigma$  where

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad (9.1)$$

and  $\Sigma_{11} \dots p_1 \times p_1$  is the covariance matrix of  $\underline{x}_1$ . Then  $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$  is known to be the canonical correlation matrix. Without loss of generality, we assume that  $p_1 \leq p_2$ .  $\rho_1^2 \geq \dots \geq \rho_{p_1}^2$  denote the eigenvalues of  $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ . Here,  $\rho_1, \dots, \rho_{p_1}$  are known as canonical correlations where  $\rho_i$  is the positive square root of  $\rho_i^2$ . Now let  $\underline{\alpha}_i$  and  $\underline{\beta}_i$  denote the eigenvectors of  $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$  and  $\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$  respectively corresponding to  $\rho_i^2$ . Then  $\underline{\alpha}_1 \underline{x}_1, \dots, \underline{\alpha}_{p_1} \underline{x}_1$  and  $\underline{\beta}_1 \underline{x}_2, \dots, \underline{\beta}_{p_1} \underline{x}_2$  are known as canonical variables. One of the important problems in the area of canonical correlation analysis is to find out the number of canonical correlations which are significantly different from zero. In this section, we discuss some procedures for testing the hypothesis on the rank of the canonical correlation matrix when the underlying distribution is multivariate normal.

Let  $X:n \times p$  be a random matrix such that  $E(X) = 0$  and  $E(X'X) = n\Sigma$ . Also let,

$$S = X'X = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \quad (9.2)$$

where  $S_{11}$  is of order  $p_1 \times p_1$ . In addition, let  $r_1^2 \geq \dots \geq r_{p_1}^2$  denote the eigenvalues of

$S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$ . Then  $r_1, \dots, r_{p_1}$  are known as the sample canonical correlations where  $r_i$  is the positive square root of  $r_i^2$ . Various functions of  $r_1^2, \dots, r_{p_1}^2$  were proposed in the literature as test statistics for determination of the rank of the canonical correlation matrix. We will review these procedures in this section.

We first assume that the rows of  $X$  are distributed independently as multivariate normal. In this case, Bartlett (1947) proposed a procedure for testing the hypothesis  $H_t$  where  $H_t$  denotes  $\rho_{t+1}^2 = \dots = \rho_{p_1}^2 = 0$ ; he also derived asymptotic distribution of the above statistic. Fujikoshi (1974) showed that the above test statistic is the LRT statistic. Hsu (1941) derived asymptotic joint distribution of the sample canonical correlations when  $H_t$  is true. When the population canonical correlations  $\rho_1, \dots, \rho_{p_1}$  have multiplicities and none of them is equal to zero, Fujikoshi (1978) derived the nonnull distribution of a single function of the sample canonical correlations whereas Krishnaiah and Lee (1979) derived asymptotic joint distribution of functions of the sample canonical correlations. The expressions derived by Krishnaiah and Lee involve multivariate normal density and multivariate Hermite polynomials. When the underlying distribution is not multivariate normal, Fang and Krishnaiah (1982) obtained results analogous to those obtained in the above paper of Krishnaiah and Lee.

Now, let the joint distribution of the elements of  $X$  be elliptically symmetric with density given by

$$f(X) = |\Sigma|^{-n/2} h(\text{tr} \Sigma^{-1} X'X) \quad (9.3)$$

Then, Krishnaiah, Lin and Wang (1985) showed that the LRT statistic for testing the hypothesis  $\rho_{t+1}^2 = \dots = \rho_{p_1}^2 = 0$  is given by

$$L(k) = \prod_{j=t+1}^{p_1} (1-r_j^2)^{n/2} \quad (9.4)$$

So, the LRT statistic is the same as when the underlying distribution is multivariate normal. They also noted that the distribution of any function of  $r_1^2, \dots, r_{p_1}^2$  is independent of the form of the underlying distribution as long as the underlying distribution belongs to the family of elliptical distributions.

We will now review some of the work reported in the literature on canonical correlation analysis when it is assumed that the observations are distributed independently and identically as elliptically symmetric with the following common density

$$f(\underline{x}) = |\Sigma|^{-1/2} h(\underline{x}' \Sigma^{-1} \underline{x}). \quad (9.5)$$

Now, let

$$c_i = \frac{\sqrt{n(r_i^2 - p_i^2)}}{2p_i^2(1-p_i^2)}. \quad (9.6)$$

Then, Murihead and Waternaux (1980) showed that  $c_1, \dots, c_{p_1}$  are asymptotically distributed independently as normal with mean 0 and variance  $(\kappa+1)$  when  $p_1^2, \dots, p_{p_1}^2$  are distinct. This is a special case of a result of Fang and Krishnaiah (1982). Krishnaiah, Lin and Wang (1985b) derived asymptotic joint distribution of the sample canonical correlations when the population canonical correlations have multiplicities and the last few population canonical correlations are zero. In particular, they showed that the joint asymptotic distribution of  $((nr_{s+1}^2/(\kappa+1)), \dots, (nr_{p_1}^2/(\kappa+1)))$  when  $H_s: p_{s+1}^2 = \dots = p_{p_1}^2 = 0$ , is the same as the joint distribution of the eigenvalues of the central Wishart matrix  $W_{p_1-s}$  with  $(p_2-s)$  degrees of freedom and  $E(W_{p_1-s}) = (p_2-s)I_{p_1-s}$ . This result is useful in implementation of certain test procedures for  $H_s$  when the sample size is large. For example, we can use  $r_{s+1}^2$  or  $(r_{s+1}^2 + \dots + r_{p_1}^2)$  as a test statistic for  $H_s$ .

We now discuss the problem of testing for the rank of the canonical correlation

matrix under correlated multivariate regression equations (CMRE) model considered by Kariya, Fujikoshi and Krishnaiah (1984). Consider the CMRE model

$$Y_i = X_i \theta_i + E_i \quad (9.7)$$

for  $i = 1, 2$ . In the above model, the rows of  $(E_1, E_2)$  are distributed independently as multivariate normal with mean vector  $0$  and covariance matrix  $\Sigma$  where

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad (9.8)$$

and  $\Sigma_{ij}$  is of order  $p_i \times p_j$ . Also,  $X_i : n \times r_i$  is the design matrix and  $\theta_i : r_i \times p_i$  is the matrix of unknown parameters for  $i = 1, 2$ . Without loss of generality, we assume that  $p_1 < p_2$ . Now, let

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \quad (9.9)$$

where  $S_{ij} = Y_i' Q Q' Y_j$  and  $Q_i = I_n - X_i (X_i' X_i)^{-1} X_i'$ . Also, let  $R = S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}$ . Kariya, Fujikoshi and Krishnaiah (1984) investigated the problem of testing the hypothesis that  $\rho_t^2 = \dots = \rho_{p_1}^2 = 0$ . They also derived the asymptotic distributions of three statistics in the null case and under local alternatives. We can test the hypothesis that  $\rho_t^2 = \dots = \rho_{p_1}^2 = 0$  by considering suitable functions of  $r_t^2, \dots, r_{p_1}^2$  like  $r_t^2, r_t^2 + \dots + r_{p_1}^2$ , etc., where  $r_1^2 \geq \dots \geq r_{p_1}^2$  are the eigenvalues of the sample canonical correlation matrix  $S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}$ .

For an application of the canonical correlation analysis in econometrics, the reader is referred to Hannan (1967) and Chow and Ray-Chowdhuri (1967).

#### 10 MODEL SELECTION METHODS FOR DETERMINATION OF THE RANK OF THE CANONICAL CORRELATION MATRIX

Let  $X' = [x_1, \dots, x_n]$   $p \times n$  be a random matrix whose columns are distributed

independently and identically as multivariate normal with common mean vector  $\underline{0}$  and covariance matrix  $\Sigma$ . Let  $\underline{x}_i$  and  $\Sigma$  be partitioned as  $\underline{x}_i = (\underline{x}_{i1}, \underline{x}_{i2})$  and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad (10.1)$$

where  $\Sigma_{ij}$  is of order  $p_i \times p_j$  and  $\underline{x}_{ij}$  is of order  $p_j \times 1$ . Let  $\rho_1^2 \geq \dots \geq \rho_s^2$  denote the first largest  $s$  eigenvalues of the population canonical correlation matrix  $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$  where  $s = \min(p_1, p_2)$ . Also, let  $r_1^2 \geq \dots \geq r_s^2$  denote the first largest  $s$  eigenvalues of  $S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}$  where

$$S = \sum_{j=1}^n \underline{x}_j \underline{x}_j' = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \quad (10.2)$$

and  $S_{ij}$  is of order  $p_i \times p_j$ . Let  $M_k (k = 0, 1, 2, \dots, s)$  denote the model for which  $\text{rank}(\Sigma_{12}) = k$ , that is, the number of nonzero canonical correlations is equal to  $k$ . Also, let  $H_k$  denote the hypothesis that  $\text{rank}(\Sigma_{12}) = k$ . Let  $L(k)$  denote the likelihood ratio test statistic for  $H_k$ . Then

$$\log L(k) = (n/2) \sum_{i=k+1}^s \log(1-r_i^2) \quad (10.3)$$

Now, let

$$C(k) = -\log L(k) + kC_n \quad (10.4)$$

where  $C_n$  satisfies the following conditions

$$(i) \lim_{n \rightarrow \infty} \{C_n/n\} = 0$$

(10.5)

$$(ii) \lim_{n \rightarrow \infty} \{C_n/\log \log n\} > p_1 p_2$$

Let  $q$  denote the true rank of  $\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1}$ . Then, Bai, Krishnaiah and Zhao (1986a) proposed to use  $\hat{q}$  as an estimate of  $q$  where  $\hat{q}$  is given by

$$G(\hat{q}) = \min\{G(0), \dots, G(s)\}. \quad (10.6)$$

The above authors also showed that  $\hat{q}$  is a strongly consistent estimate of  $q$ . Now, assume that the assumption of normality is violated but  $\underline{x}_1, \dots, \underline{x}_n$  are i.i.d. vectors with  $E(\underline{x}_1) = 0$ ,  $E(\underline{x}_1 \underline{x}_1') = \Sigma$  and  $E(\underline{x}_1' \underline{x}_1)^2 < \infty$ . When the assumption of normality is violated,  $L(k)$  need not be the LRT statistic for  $H_k$  but we can still use it in (10.4). Then Bai, Krishnaiah and Zhao (1986) showed that  $\hat{q}$  in (10.6) is still a strongly consistent estimate of  $q$  under certain conditions.

## 11. SELECTION OF ORIGINAL VARIABLES IN CANONICAL CORRELATION ANALYSIS

Let us consider the set  $(\underline{x}_1', \underline{x}_2')$  of  $p_1 + p_2$  variables. We wish to select a set of  $r_2$  important variables from the  $\underline{x}_2$  set on the basis of the degree of dependence with  $\underline{x}_1$  set.

There are  $\binom{p_2}{r_2}$  sets. Let these sets be denoted by  $\underline{x}_{f_i}$  and let the sample canonical correlation matrix between  $\underline{x}_1$  set and  $\underline{x}_{f_i}$  set be denoted by  $S_{11}^{-1} S_{1f_i} S_{ff_i}^{-1} S_{f_i 1}$ . We use the largest root of the canonical correlation matrix as a criterion to select the variables. We declare that none of these sets are important if

$$\max_i c_L(S_{11}^{-1} S_{1f_i} S_{ff_i}^{-1} S_{f_i 1}) \leq c_\alpha$$

where  $c_L(A)$  denotes the largest eigenvalue of  $A$ . If

$$\max_i c_L(S_{11}^{-1} S_{1f_i} S_{ff_i}^{-1} S_{f_i 1}) > c_\alpha$$

then the set corresponding to  $\max_i c_L(S_{11}^{-1}S_{1f}, S_{ff}^{-1}S_{f1})$  is declared to be the most important. The critical value  $c_\alpha$  is chosen such that

$$P[\max_i c_L(S_{11}^{-1}S_{1f}, S_{ff}^{-1}S_{f1}) \leq c_\alpha | H] = (1-\alpha)$$

and  $H: \Sigma_{12} = 0$ . In other words, the critical value  $c_\alpha$  is chosen such that the probability of declaring that none of the  $\binom{p_2}{r_2}$  sets are important when in fact none of the variables in the  $x_2$  set are correlated with  $x_1$  set. But the distribution of  $\max_i c_L(S_{11}^{-1}S_{1f}, S_{ff}^{-1}S_{f1})$  is very complicated to derive. So, we use the following bound to get an approximate value of  $c_\alpha$ :

$$\begin{aligned} P[c_L(S_{11}^{-1}S_{12}, S_{22}^{-1}S_{21}) \leq c_\alpha | H] \\ \leq P[\max_i c_L(S_{11}^{-1}S_{1f}, S_{ff}^{-1}S_{f1}) \leq c_\alpha | H] = (1-\alpha) \end{aligned}$$

We will now discuss an alternative procedure for the selection of the best subset of  $q$  variables from the  $x_2$  set and let  $x_{f_i}$  ( $i = 1, 2, \dots, \binom{p_2}{q}$ ) denote a subset of  $q$  variables from the  $p_2$  variables  $x_2$ . As before, let  $\Sigma_{11}^{-1}\Sigma_{1f_i}, \Sigma_{f_i f_i}^{-1}\Sigma_{f_i 1}$  denote the canonical correlation matrix connected with  $x_1$  set and  $x_{f_i}$  set. Let  $\psi_i$  denote a suitable function of the eigenvalues of the above matrix. Also, let  $\hat{\psi}_i$  denote the corresponding function of the eigenvalues of  $S_{11}^{-1}S_{1f_i}, S_{f_i f_i}^{-1}S_{f_i 1}$ . In addition, let  $\psi_1, \dots, \psi_{p_0}$  be ordered as  $\psi_{[1]} \geq \psi_{[2]} \geq \dots \geq \psi_{[p_0]}$  where  $p_0 = \binom{p_2}{q}$ . Then, the subset associated with the maximum of  $\hat{\psi}_1, \dots, \hat{\psi}_{p_0}$  is declared to be the best subset. Suppose  $\hat{\psi}_1$  is the largest of  $\hat{\psi}_i$ s. In this case, the

probability of correct decision is given by the probability of  $\hat{\psi}_i$  being greater than  $\hat{\psi}_j$  ( $j = 1, \dots, i-1, i+1, \dots, p_0$ ) when  $\psi_i$  is greater than  $\psi_j$  for  $j \neq i$ . This probability involves nuisance parameters. One may use bounds which are free from nuisance parameters.

We now discuss the problem of studying the effect of additional variables on the canonical correlations. Consider two sets of variables  $\underline{x}_1 : p_1 \times 1$  and  $\underline{y}_1 : q_1 \times 1$ . Without loss of generality, we assume that  $p_1 \leq q_1$ . Suppose the sets of variables  $\underline{x}_1$  and  $\underline{y}_1$  are augmented to  $\underline{x} : p \times 1$  and  $\underline{y} : q \times 1$  by adding extra sets of variables  $\underline{x}_2 : p_2 \times 1$  and  $\underline{y}_2 : q_2 \times 1$  respectively. Also, we assume that  $(\underline{x}, \underline{y})$  is distributed as multivariate normal with mean vector  $\underline{\mu}$  and covariance matrix  $\Sigma$  where

$$\Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}$$

and  $\Sigma_{xx}$  is the covariance matrix of  $\underline{x}$ . Let  $\rho_1 \geq \dots \geq \rho_{p_1}$  denote the canonical correlations between the sets  $\underline{x}_1$  and  $\underline{y}_1$  and let  $\tilde{\rho}_1 \geq \dots \geq \tilde{\rho}_p$  denote the canonical correlations between  $\underline{x}$  and  $\underline{y}$ . Also, let  $\delta_\alpha = \tilde{\rho}_\alpha - \rho_\alpha$  for  $\alpha = 1, 2, \dots, p_1$ . Then  $\delta_\alpha > 0$ . Next, let

$$S = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix}$$

denote the sample covariance matrix based on  $(n+1)$  observations on  $(\underline{x}, \underline{y})$  and the sample canonical correlations  $\tilde{r}_1 \leq \dots \leq \tilde{r}_p$  are the positive square roots of the eigenvalues of  $S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx}$ . Similarly, let  $r_1 \geq \dots \geq r_{p_1}$  denote the sample canonical correlations based on  $(n+1)$  observations on  $(\underline{x}_1, \underline{y}_1)$ .

Now, let  $f(d_1, \dots, d_{p_1})$  be a continuously differentiable function in a neighborhood of  $\underline{d} = \underline{\delta}$  where  $\underline{d} = (d_1, \dots, d_{p_1})$  and  $\underline{\delta} = (\delta_1, \dots, \delta_{p_1})$ . Then, Fujikoshi, Krishnaiah and Schmidhammer (1985) showed that  $\sqrt{n}\{f(d_1, \dots, d_{p_1}) - f(\delta_1, \dots, \delta_{p_1})\}$  is distributed asymptotically as normal with

mean zero and certain variance  $\sigma^2$ . When  $f(d_1, \dots, d_{p_1}) = d_1$ , and  $p_1 = p$  or  $q_1 = q$  the above result was derived by Wijsman (1986). The result of Fujikoshi, Krishnaiah and Schmidhammer (1985) can be used to find out whether the addition of new variables to one or both of the sets  $\underline{x}_1$  and  $\underline{y}_1$  will have effect on functions of the canonical correlations. For example, we can draw inference as to whether the addition of variables will increase the values of the largest canonical correlation, sum of the canonical correlations, etc. If there is no significant increase, we will conclude that the new variables are not important in explaining the association between the two sets of variables. Fujikoshi (1985) proposed a procedure based on an information theoretic criterion to select best variables in canonical correlation analysis.

## 12. REDUCTION OF DIMENSIONALITY AND THE STRUCTURE OF DEPENDENCE IN TWO-WAY CONTINGENCY TABLE

Consider two-way contingency table and let  $p_{ij}$  ( $i = 1, 2, \dots, r+1$ ;  $j = 1, 2, \dots, s+1$ ) denote the probability of an observation falling in  $i$ -th row and  $j$ -th column. We will consider the model

$$p_{ij} = p_{i.} p_{.j} \zeta_{ij} \quad (12.1)$$

where  $p_{i.} = p_{i1} + \dots + p_{i,s+1}$ ,  $p_{.j} = p_{1j} + \dots + p_{r+1,j}$  and  $\zeta_{ij}$  is an unknown constant. In a number of situations, we are interested in studying the structure of dependence between rows and columns if  $p_{ij} \neq p_{i.} p_{.j}$ . If we know the structure of dependence, we can estimate the unknown parameters more efficiently. Now, let  $F = (f_{ij})$  where  $f_{ij} = p_{ij} / \sqrt{p_{i.} p_{.j}}$ . From the singular value decomposition of the matrix, it is known (e.g., see Lancaster (1969)) that

$$F = \xi_0^* \eta_0^* \delta_0 + \sum_{u=1}^T \delta_u \xi_u^* \eta_u^* \quad (12.2)$$

where  $\delta_0 \geq \dots \geq \delta_r$  are the eigenvalues of  $F$ ,  $\xi_u^*$  is the eigenvector of  $FF'$  corresponding to  $\delta_u^2$  and  $\eta_u^*$  is the eigenvector of  $F'F$  corresponding to  $\delta_u^2$ . Here  $\delta_0 = 1$ ,  $\xi_0^* =$

$(\sqrt{p_{i1}}, \sqrt{p_{i2}}, \dots, \sqrt{p_{is}})$  and  $\pi_{i0}^* = (\sqrt{p_{i1}}, \sqrt{p_{i2}}, \dots, \sqrt{p_{is}})$ . We will now review the work of O'Neill (1978a, 1978b, 1980) and Bhaskara Rao, Krishnaiah and Subramanyam (1985) for testing for the rank of the matrix  $\zeta$ . We also review the work of Bai, Zhao and Krishnaiah (1986) for determination of the rank of  $\zeta$  by using model selection methods. Without loss of generality we assume that  $r \leq s$  in the sequel.

Let  $n_{ij}$  denote the frequency in  $i$ -th row and  $j$ -th column,  $n_{i.} = n_{i1} + \dots + n_{i,s+1}$  and  $n_{.j} = n_{1j} + \dots + n_{r+1,j}$ . Also, let  $B = (b_{ij})$  where  $b_{ij} = n_{ij} / \sqrt{n_{i.} n_{.j}}$ . Now, let  $\delta_0^2 \geq \dots \geq \delta_r^2$  denote the eigenvalues of  $BB'$  where  $\delta_0^2 = 1$ . We assume that  $n$  is fixed and the joint distribution of the cell frequencies is given by

$$n! \prod_{i,j} \frac{1}{n_{ij}!} p_{ij}^{n_{ij}} \quad (12.3)$$

The classical test statistic for testing the hypothesis  $p_{ij} = p_{i.} p_{.j}$  of independence is given by

$$\chi_0^2 = \sum_{i=1}^{r+1} \sum_{j=1}^{s+1} (n_{ij} - (n_{i.} n_{.j} / n))^2 / n_{i.} n_{.j} \quad (12.4)$$

When the null hypothesis is true,  $\chi_0^2$  is distributed asymptotically as chi-square with  $rs$  degrees of freedom. The above hypothesis is equivalent to testing the hypothesis that  $\rho_1^2 = \dots = \rho_r^2 = 0$  and it can be used by using  $\hat{\rho}_1^2 + \dots + \hat{\rho}_r^2$  as a test statistic. This test is equivalent to the chi-square test for independence since  $\chi_0^2 = n(\hat{\rho}_1^2 + \dots + \hat{\rho}_r^2)$ . Now, let  $H_t$  denote the hypothesis that  $\rho_t^2 = 0$ . This hypothesis is equivalent to the hypothesis that the rank of  $\zeta$  is  $t$ . O'Neil ((1978a),(1978b)) showed that the joint asymptotic distribution of  $n\hat{\rho}_1^2, \dots, n\hat{\rho}_r^2$  when  $H_t$  is true, is the same as the joint distribution of the eigenvalues of the central Wishart matrix  $W$  with  $s$  degrees of freedom and  $E(W) = s \mathbf{I}_r$ . Tables for percentage points of the largest eigenvalue of the central Wishart matrix are given in Krishnaiah (1980).

We will now review the work of Bhaskara Rao, Krishnaiah and Subramanyam (1985) for determination of the rank of  $\zeta$ . They suggested functions of  $\hat{p}_1^2, \dots, \hat{p}_r^2$  as test statistics for testing  $H_1$ . For example, one may use  $\hat{p}_1^2, \hat{p}_1^2 + \dots + \hat{p}_r^2$  as test statistics. The above authors also suggested the following simultaneous test procedure. We accept or reject  $H_1$  according as

$$\hat{p}_1^2 \leq c_\alpha \quad (12.5)$$

where

$$P[\hat{p}_1^2 \leq c_\alpha | H_1] = (1-\alpha). \quad (12.6)$$

If  $H_1$  is accepted and  $H_{t-1}$  is rejected, then the rank of  $\zeta$  is  $t$ . Bhaskara Rao, Krishnaiah and Subramanyam (1985) derived asymptotic joint distribution of functions of  $\hat{p}_1^2, \dots, \hat{p}_r^2$  when  $p_1^2, \dots, p_r^2$  have multiplicities. O'Neil (1978a) suggested using  $n(\hat{p}_t^2 + \dots + \hat{p}_r^2)$  as a test statistic for testing the hypothesis that the rank of  $\zeta$  is  $t$ . In general, we can use a suitable function of  $\hat{p}_t^2, \dots, \hat{p}_r^2$  like the above test statistic or  $n\hat{p}_t^2$  to test the hypothesis that the rank of  $p$  is  $t$ . But, unfortunately, the distributions of the above test statistics involve nuisance parameters even asymptotically. As an ad hoc procedure, one can replace the nuisance parameters with their consistent estimates.

Bai, Krishnaiah and Zhao (1986b) proposed the following procedure for determination of the rank of  $P = (p_{ij})$ . Let

$$G(k) = n \sum_{j=k+1}^r \hat{p}_j^2 + kC_n \quad (12.7)$$

where  $C_n$  satisfies the following conditions:

$$(i) \quad \lim_{n \rightarrow \infty} (C_n/n) = 0$$

(12.8)

$$(ii) \quad \lim_{n \rightarrow \infty} (C_n/\log \log n) = \infty.$$

Then, the unknown rank  $q$  of  $P$  is estimated with  $\hat{q}$  where  $\hat{q}$  is given by

$$G(\hat{q}) = \min \{G(1), \dots, G(r)\}.$$

Bai, Krishnaiah and Zhao (1986b) showed that  $\hat{q}$  is a consistent estimate of  $q$ .

## REFERENCES

1. Akaike, H. (1972). Information theory and an extension of the maximum likelihood principle. In Proceedings of the Second International Symposium on Information Theory, Suppl. to Problems of Control and Information Theory, pp. 267-281.
2. Amemiya, Y. and Fuller, W. (1984). Estimation for the multivariate errors-in-variables model with estimated error covariance matrix. Ann. Statist. 12 497-509.
3. Anderson, T.W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. Ann. Math. Statist. 22 327-351 (correction, Ann. Statist. 8 (1980), 1400).
4. Anderson, T.W., (1984). Estimating linear statistical relationships. Ann. Statist. 12 1-45.
5. Anderson, T.W., (1985). Components of variance in MANOVA. In Multivariate Analysis-VI (P.R. Krishnaiah, editor), pp. 1-8. North-Holland Publishing Company.
6. Anderson, B.M., Anderson, T.W. and Olkin, I. (1986). Maximum likelihood estimators and likelihood ratio criteria in multivariate components of variance. To appear in Ann. Statist.
7. Bai, Z.D. (1984). A note on asymptotic joint distribution of the eigenvalues of a noncentral multivariate F matrix. Technical Report No. 84-49. Center for Multivariate Analysis, University of Pittsburgh.
8. Bai, Z.D., Krishnaiah, P.R. and Liang, W.Q. (1984). On asymptotic joint distribution of the eigenvalues of a noncentral multivariate F matrix. To appear in Sankhyā.
9. Bai, Z.D., Krishnaiah, P.R. and Zhao, L.C., (1986a). Inference on the ranks of regression matrix and canonical correlation matrix using model selection methods. Technical Report No. 86- . Center for Multivariate Analysis, University of Pittsburgh.
10. Bai, Z.D., Krishnaiah, P.R. and Zhao, L.C. (1986b). Model selection methods for log linear models. Technical Report No. 86- . Center for Multivariate Analysis, University of Pittsburgh.
11. Bartlett, M.S., (1948). A note on the statistical estimation of demand and supply relations from time series. Econometrica , 16 323-329.

12. Bhaskar Rao, M., Krishnaiah, P.R. and Subramanyam, K. (1985). A structure theorem on bivariate positive quadrant dependent distributions and tests for independence in two-way contingency tables. Technical Report No. 85-48, Center for Multivariate Analysis, University of Pittsburgh.
13. Chow, G.C. and Ray-Choudhuri, D.K. (1967). An alternative proof of Hannan's theorem on canonical conditions and multiple equation systems. Econometrica, 35, 139-142.
14. Corsten, L.C.A. and van Eijnsbergen, A.C. (1972). Multiplicative effects in two-way analysis of variance. Statistica Neerlandica, 26 61-68.
15. Fang, C. and Krishnaiah, P.R., (1982). Asymptotic distributions of functions of the eigenvalues of some random matrices for non-normal populations. J. Multivariate Anal., 12 39-63.
16. Fisher, R.A. and Mackenzie, W.A. (1923). Studies in crop variation II. The manurial response of different potato varieties. J. of Agricultural Science, 13 311-
17. Fisher, R.A., (1939). The sampling distribution of some statistics obtained from nonlinear equations. Ann. Eugen., 9, 283-349.
18. Fujikoshi, Y. (1974). The likelihood ratio tests for the dimensionality of regression coefficients. J. Multivariate Anal. 4, pp. 327-340.
19. Fujikoshi, Y. (1977). Asymptotic expansions for the distributions of some multivariate tests. In Multivariate Analysis IV, (P.R. Krishnaiah, editor), pp. 55-71. North-Holland Publishing Company.
20. Fujikoshi, Y. (1978). Asymptotic expansions for the distributions of some functions of the latent roots of matrices in three situations. J. Multivariate Anal., 8 63-72.
21. Fujikoshi, Y. (1985). Selection of variables in discriminant analysis and canonical correlation analysis. In Multivariate Analysis VI, (P.R. Krishnaiah, editor). North-Holland Publishing Company.
22. Fujikoshi Y., Krishnaiah, P.R. and Schmidhammer, J. (1985). Effect of additional variables in principal component analysis, discriminant analysis and canonical correlation analysis. Technical Report No. 85-31, Center for Multivariate Analysis, University of Pittsburgh.
23. Gollob, H.F. (1968). A statistical model which combines features of factor analysis and analysis of variance techniques, Psychometrika, 33 73-116.

24. Goodman, N.R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). Ann. Math. Stat., 34 152-176.
25. Hannan, E.J., (1967). Canonical correlation and multiple equation systems in economics. Econometrica, 12 124-138.
26. Hotelling, H.O. (1935). The most predictable criterion. J. of Educational Psychology, 26 139-142.
27. Hotelling, H.O. (1936). Relations between two sets of variates. Biometrika, 28 321-377.
28. Hsu, P.L. (1941a). On the limiting distribution of roots of a determinantal equation. J. London Math. Soc., 16 183-194.
29. Hsu, P.L. (1941b). On the problem of rank and the limiting distribution of Fisher's test function. Ann. Eugen., 11 39-41.
30. Izenman, A.J., (1975). Reduced-rank regression for the multivariate linear model. J. Multivariate Anal., 5, 248-264.
31. Johnson, D.E. and Graybill, F.A. (1972). An analysis of a two-way model with interaction and no replication. J. Amer. Statist. Assoc., 67 862-868
32. Kariya, T., Fujikoshi, Y. and Krishnaiah, P.R. (1984). Tests for independence of two multivariate regression equations with different design matrices. J. Multivariate Anal., 15 383-407
33. Kelker, D., (1970). Distribution theory of spherical distribution and a location scale parameter generalization. Sankhyā A, 43 419-430
34. Krishnaiah, P.R. (1965). Multiple comparison tests in multi-response experiments Sankhyā A, 27 65-72
35. Krishnaiah, P.R. (1969). Simultaneous test procedures under general MANOVA models. In Multivariate Analysis II, (P.R. Krishnaiah, editor) Academic Press, Inc
36. Krishnaiah, P.R. and Armitage, T.V. (1970). On a multivariate F distribution. In Essays in Probability and Statistics (R.C. Bose et al. eds) University of North Carolina, Chapel Hill
37. Krishnaiah, P.R. and Waikar, V.B. (1971). Simultaneous tests for equality of latent roots against certain alternatives I. Ann. Inst. Statist. Math., 23 451-468

38. Krishnaiah, P.R. and Waikar, V.B. (1972). Simultaneous test for equality of latent roots against certain alternatives-II. Ann. Inst. Statist. Math., 24 81-85.
39. Krishnaiah, P.R. and Schuurmann, F.J. (1974). On the evaluation of some distributions that arise in simultaneous tests for the equality of the latent roots of the covariance matrix. J. Multivariate Anal., 4, 265-282.
40. Krishnaiah, P.R. (1976). Some recent developments on complex multivariate distributions. J. Multivariate Anal., 6, 1-30.
41. Krishnaiah, P.R. and Lee, J.C. (1979). On the asymptotic joint distributions of certain functions of the eigenvalues of some random matrices. J. Multivariate Anal., 9 248-258.
42. Krishnaiah, P.R. (1980). Computations of some multivariate distributions. In Handbook of Statistics, Vol. 1 (P.R. Krishnaiah, editor). North-Holland Publishing Company, pp. 745-972.
43. Krishnaiah, P.R. and Yochmowitz, M.G. (1980). Inference on the structure of interaction in two-way classification model. In Handbook of Statistics, Vol. 1: Analysis of Variance (P.R. Krishnaiah, editor) pp. 973- 994.
44. Krishnaiah, P.R. (1982). Selection of variables under univariate regression models. In Handbook of Statistics, Vol. 2 Classification, Pattern Recognition and Reduction of Dimensionality (P.R. Krishnaiah, editor), North-Holland Publishing Company, pp. 805-820.
45. Krishnaiah, P.R. and Lin, J., (1984). Complex elliptical distributions. To appear in Communications in Statistics.
46. Krishnaiah, P.R., Lin, J. and Wang, L. (1985a). Tests for the dimensionality of the regression matrices when the underlying distributions are elliptically symmetric. Technical Report No. 85-36. Center for Multivariate Analysis, University of Pittsburgh.
47. Krishnaiah, P.R., Lin, J. and Wang, L., (1985b). Inference on the ranks of the canonical correlation matrices for elliptically symmetric populations. Technical Report No. 85-14. Center for Multivariate Analysis, University of Pittsburgh.
48. Lancaster, H.O. (1969). The Chi-Square Distribution. John Wiley, New York.
49. Lawley, D.N. (1956). Tests of significance for the latent roots of covariance and correlation matrices. Biometrika, 43 128-136.
50. Mandel, J. (1969). Partitioning the interaction in analysis of variance. J. Res. Nat. Bur. Standards, Sect. B 73B, 309-328.

- 51 Muirhead, R.J. and Waternaux, C. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. Biometrika, 67, 31-43.
- 52 Moustafa, M.D. (1957). Tests of hypotheses on a multivariate population, some of the variables being continuous and the rest categorical. Mimeo Series No 179, Institute of Statistics, University of North Carolina.
- 53 Olkin, I. and Tate, R.F. (1961). Multivariate correlation models with mixed discrete and continuous variables. Ann. Math. Statist., 32448-465.
- 54 O'Neill, M.E. (1978a). Asymptotic distributions of the canonical correlations from contingency tables. Austral. J. Statist., 20, 75-82.
- 55 O'Neill, M.E. (1978b). Distributional expansions for canonical correlations from contingency tables. J. Roy. Statist. Soc. Ser. B, 40, 303-312.
- 56 O'Neill, M.E. (1980). A note on canonical correlations from contingency tables. Austral. J. Statist.
- 57 Rao, C.R. (1948). Tests of significance in multivariate analysis. Biometrika, 35, 58-79.
- 58 Rao, C.R., (1973). Linear Statistical Inference and Its Applications. John Wiley & Sons.
- 59 Rao, C.R., (1983). Likelihood ratio tests for relationships between two covariance matrices. In Studies in Econometrics, Time Series and Multivariate Statistics (Amemiya, T., Karlin, S. and Goodman, L., editors), pp. , Academic Press, Inc., New York.
- 60 Rao, C.R. (1985). Tests for dimensionality and interactions of mean vectors under general and reducible covariance structures. J. Multivariate Anal., 16, 173-184.
- 61 Rissanen, J., (1973). Modeling by shortest data description. Automatica, 14, 465-471.
- 62 Roy, J. (1958). Step-down procedure in multivariate analysis. Ann. Math. Statist., 29, 1177-1187.
- 63 Schmidhammer, J.L. (1982). On the selection of variables under regression models using Krishnaiah's finite intersection tests. In Handbook of Statistics, Vol. 2: Classification, Pattern Recognition, and Reduction of Dimensionality (P.R. Krishnaiah and L.N. Kanal, editors). North-Holland Publishing Company.

64. Schott, J.R. and Saw, J.G. (1984). A multivariate one-way classification model with random effects. J. Multivariate Anal., 14 1-12.
65. Schuurmann, F.J., Krishnaiah, P.R. and Chattopadhyay, A.K. (1973). On the distributions of the ratios of the extreme roots to the trace of the Wishart matrix. J. Multivariate Anal., 3 445-453.
66. Schwartz, G. (1978). Estimating the dimension of a model. Ann. Statist., 6, 461-464.
67. Siotani, M. (1959). The extreme value of the generalized distances of the individual points in the multivariate normal sample. Ann. Inst. Statist. Math., 10 183-203.
68. Siotani, M. (1960). Notes on multivariate confidence bounds. Ann. Inst. Statist. Math., 11 167-182.
69. Siotani, M. (1961). The extreme value of the generalized distances and its applications. Bulletin of the International Statistical Institute, 38 591-599.
70. Thompson, M.L. (1978a). Selection of variables in multiple regression; Part I. A review and evaluation. Int. Statist. Rev., 46, 1-19.
71. Thompson, M.L. (1978b). Selection of variables in multiple regression; Part II. Chosen procedures, computations and examples. Int. Statist. Rev., 46 129-146.
72. Tintner, G. (1945). A note on rank, multicollinearity and multiple regression. Ann. Math. Statist., 16 304-308.
73. Tukey, J.W. (1949). One degree of freedom for non-additivity. Biometrics, 5, 232-242.
74. Wijsman, R.A. (1986). Asymptotic distribution of the increase of the largest canonical correlation when one of the vectors is augmented. J. Multivariate Anal.
75. Williams, E.J. (1952). The interpretation of interactions in factorial experiments. Biometrika, 39 65-81.
76. Wooding, R.A. (1956). The multivariate distribution of complex normal variables. Biometrika, 43 212-215.
77. Yochmowitz, M.G. and Cornell, R.G. (1978). Stepwise tests for multiplicative components of interaction. Technometrics, 20 79-84.

78. Zhao, L.C., Krishnaiah, P.R. and Bai, Z.D., (1985a). On detection of number of signals in presence of white noise. Technical Report No. 85-37. Center for Multivariate Analysis, University of Pittsburgh.
79. Zhao, L.C., Krishnaiah, P.R. and Bai, Z.D. (1985b). On detection of number of signals in presence of colored noise using information theoretic criteria. Technical Report No. 85-39. Center for Multivariate Analysis, University of Pittsburgh.

END

DT/C

8-86